

OkTrends: Dating Research from OkCupid

[Join us!](#)

We Experiment On Human Beings!

July 28th, 2014 by [Christian Rudder](#)

I'm the first to admit it: we might be popular, we might create a lot of great relationships, we might blah blah blah. But OkCupid doesn't really know what it's doing. Neither does any other website. It's not like people have been building these things for very long, or you can go look up a blueprint or something. Most ideas are bad. Even good ideas could be better. Experiments are how you sort all this out. Like this young buck, trying to get a potato to cry.



We noticed recently that people didn't like it when Facebook “experimented” with their news feed. Even the FTC is [getting involved](#). But guess what, everybody: if you use the Internet, you're the subject of hundreds of experiments at any given time, on every site. That's how websites work.

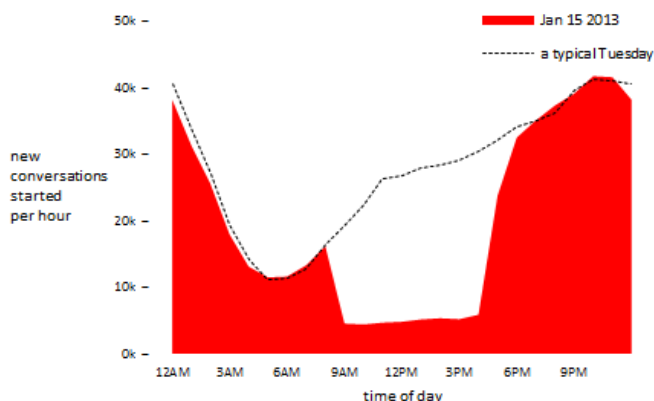
Here are a few of the more interesting experiments OkCupid has run.

Experiment 1: LOVE IS BLIND, OR SHOULD BE

OkCupid's ten-year history has been the epitome of the old saying: two steps forward, one total fiasco. A while ago, we had the genius idea of an app that set up blind dates; we spent a year and a half on it, and it was gone from the app store in six months.

Of course, *being geniuses*, we chose to celebrate the app's release by removing all the pictures from OkCupid on launch day. “Love Is Blind Day” on OkCupid—January 15, 2013.

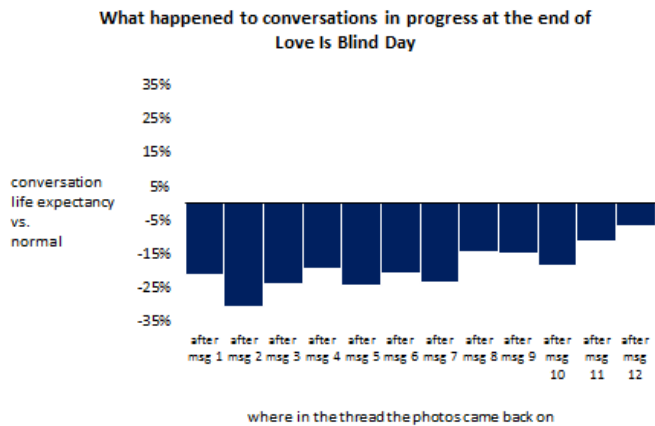
All our site metrics were way down during the “celebration”, for example:



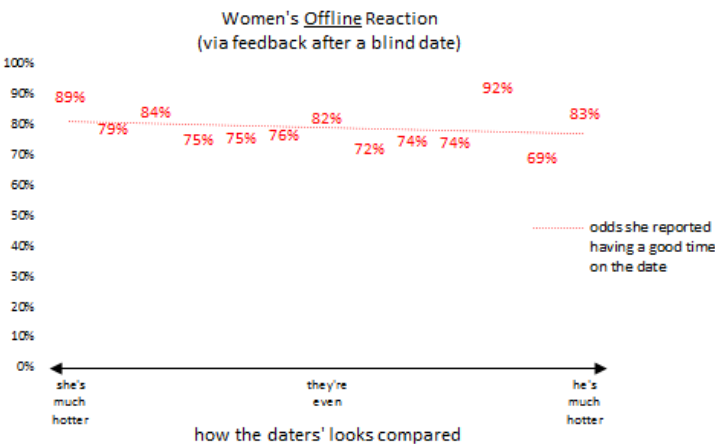
But by comparing Love Is Blind Day to a normal Tuesday, we learned some very interesting things. In those 7 hours without photos:

- people responded to first messages **44% more often**
- conversations went deeper
- contact details (phone numbers, emails) were exchanged more quickly
- in short, OkCupid worked better

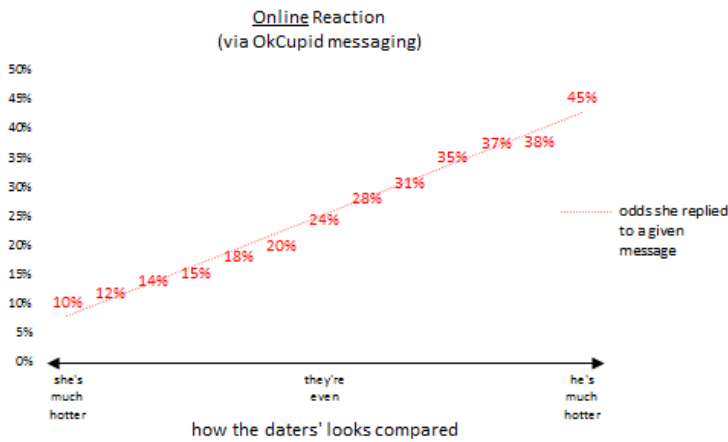
And it wasn't that "looks weren't important" to the users who'd chosen to stick around. When the photos were restored at 4PM, 2,200 people were in the middle of conversations that had started "blind". Those conversations melted away. The goodness was gone, in fact worse than gone. It was like we'd turned on the bright lights at the bar at midnight.



This whole episode made me curious, so I went and looked up the data for the people who had *actually used* the blind date app. I found a similar thing: once they got to the date, they had a good time more or less **regardless of how good-looking their partner was**. Here's the female side of the experience (the male is very similar).



Oddly, it appears that having a better-looking blind date made women slightly *less* happy —my operating theory is that hotter guys were assholes more often. Anyhow, the fascinating thing is the *online* reaction of those exact same women was just as judgmental as everyone else's:

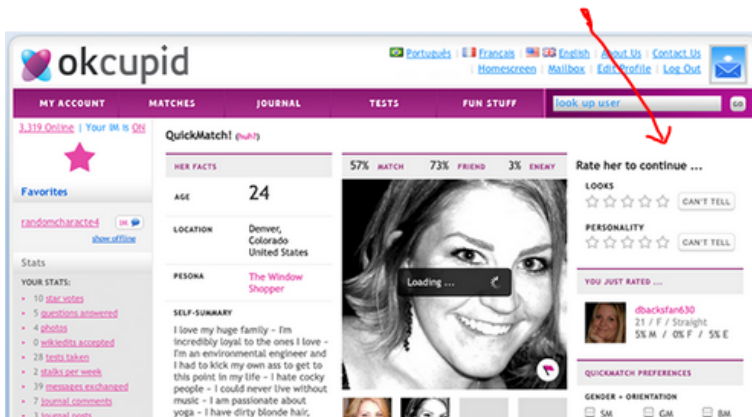


Basically, people are exactly as shallow as their technology allows them to be.

Experiment 2: SO WHAT'S A PICTURE WORTH?

All dating sites let users rate profiles, and OkCupid's original system gave people two separate scales for judging each other, "personality" and "looks."

I found this old screenshot. The "loading" icon over the picture pretty much sums up our first four years. Anyhow, here's the vote system:



Our thinking was that a person might not be classically gorgeous or handsome but *could* still be cool, and we wanted to recognize that, which just goes to show that when OkCupid started out, the only thing with more bugs than our HTML was our understanding of human nature.

Here's some data I dug up from the backup tapes. Each dot here is a person. The two scores are within a half point of each other for 92% of the sample after just 25 votes (and that percentage approaches 100% as vote totals get higher).

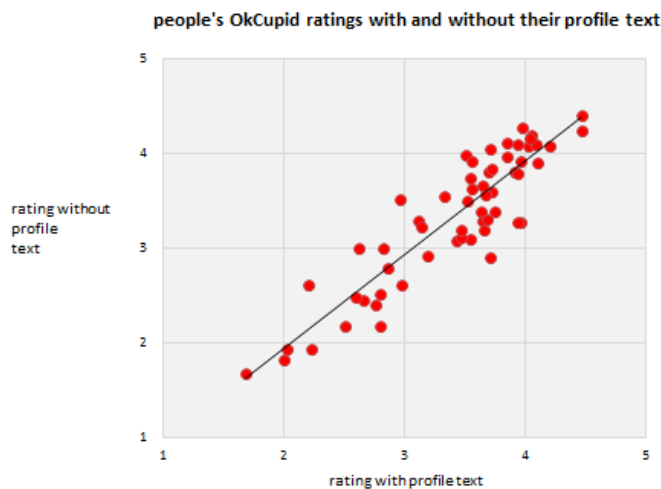


In short, according to our users, "looks" and "personality" were the same thing, which of course makes perfect sense because, you know, this young female account holder, with a 99th percentile personality:



...and whose profile, by the way, *contained no text*, is just so obviously a really cool person to hang out and talk to and clutch driftwood with.

After we got rid of the two scales, and replaced it with just one, we ran a direct experiment to confirm our hunch—that people just look at the picture. We took a small sample of users and half the time we showed them, we hid their profile text. That generated two independent sets of scores for each profile, one score for “the picture and the text together” and one for “the picture alone.” Here’s how they compare. Again, each dot is a user. Essentially, the text is less than 10% of what people think of you.



So, your picture *is* worth that fabled thousand words, but your actual words are worth...almost nothing.

Experiment 3: THE POWER OF SUGGESTION

The ultimate question at OkCupid is, does this thing even work? By all our internal measures, the “match percentage” we calculate for users is very good at predicting relationships. It correlates with message success, conversation length, whether people actually exchange contact information, and so on. But in the back of our minds, there’s always been the possibility: maybe it works just because we tell people it does. Maybe people just like each other because they think they’re supposed to? Like how Jay-Z still sells albums?

To test this, we took pairs of bad matches (actual 30% match) and told them they were exceptionally good for each other (displaying a 90% match.)† Not surprisingly, the users sent more first messages when we said they were compatible. After all, that’s what the site teaches you to do.

For users whose actual match percentage is 30%... when we DISPLAYED			
	30% match	60% match	90% match
Odds of sending one message:	14.2%	16.5%	16.9%

But we took the analysis one step deeper. We asked: does the displayed match percentage cause more than just that first message—does the mere suggestion cause people to actually *like* each other? As far as we can measure, yes, it does.

When we *tell* people they are a good match, they act as if they are. Even when they should be wrong for each other.

For users whose actual match percentage is 30%... when we DISPLAYED			
	30% match	60% match	90% match
Odds of sending one message:	14.2%	16.5%	16.9%
Given that they sent one message, odds of exchanging four :	9.7%	16.0%	17.4%

The four-message threshold is our internal measure for a real conversation. And though the data is noisier, this same “higher display means more success” pattern seems to hold when you look at contact information exchanges, too.

This got us worried—maybe our matching algorithm was just garbage and it’s only the power of suggestion that brings people together. So we tested things the other way, too: we told people who were actually *good* for each other, that they were bad, and watched what happened.

Here’s the whole scope of results (I’m using the odds of exchanging four messages number here):

Odds of a single message turning into a conversation				
		number DISPLAYED to them		
		30% match	60% match	90% match
ACTUAL compatibility of users	30% match	10%	16%	17%
	60% match	13%	13%	16%
	90% match	16%	17%	20%

As you can see, the ideal situation is the lower right: to both be *told* you’re a good match, and at the same time actually *be* one. OkCupid definitely works, but that’s not the whole story. And if you have to choose only one or the other, the mere myth of compatibility works just as well as the truth. Thus the career of someone like Doctor Oz, in a nutshell. And, of course, to some degree, mine.

If you like the this blog, you'll love this book: *Dataclysm*. I spent the last three years writing it. You should [buy it here!](#)

Journalists, please write press@okcupid.com for more information.

Leave a Reply

Name

Email

Comment