

Author notes

Spouses' faces are similar but do not become more similar with time

[Pin Pin Tea-makorn](mailto:pinnaree@stanford.edu) (pinnaree@stanford.edu) and [Michal Kosinski](mailto:michalk@stanford.edu) (michalk@stanford.edu)

Stanford University, Graduate School of Business

Article appeared in **Scientific Reports (by Nature)** and is freely available at www.nature.com/articles/s41598-020-73971-8

Summary:

Both old wives' tales and psychological literature posit that spouses' faces become more similar over time. Scholars have argued that partners tend to occupy the same environments, engage in the same activities, eat the same food, and mimic each other's emotions—and as these factors can also influence facial appearance, their faces should converge with time. For example, if the partners smile a lot—and make each other smile—they should co-develop similar smile lines.

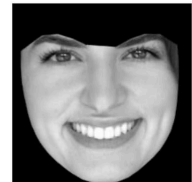
Surprisingly, this widely accepted belief is supported by a [single 1987 study](#) of 12 couples, whose results have never been replicated. In their paper, [Pin Pin Tea-makorn](#) and [Michal Kosinski](#), of Stanford's Graduate School of Business, examined this widely held belief in a large sample of 517 couples. Using facial images taken at the beginning of the marriage and 20 to 69 years later, as well as two measures of facial similarity (human judgment and a state-of-the-art facial recognition algorithm), they showed that while spouses' faces tend to be similar at the beginning of marriage, they do not converge over time.

Figure 1 shows the task presented to both human judges and the facial recognition algorithm. Detailed results are presented in Figures 2 and 3 at the end of this document.

Rank Faces Instructions (Click to collapse)

For this set of photographs, please write a number (1 - 6) in each of the blanks underneath the boxes. Put a 1 under the photograph of the person who **MOST CLOSELY RESEMBLES** the person in the target photo, a 2 under the photo of the person who is the next closest resemblance), and so on.

Note Each photo has to have a different ranking, i.e., a ranking 1, 3, 4, 2, 2, 6 is not acceptable.



Submit

Figure 1: An example stimulus set shown to both the facial recognition algorithm and human judges. It is analogous to one used in the original [1987 study](#). (To protect participants' privacy, we used photos of our colleagues.)

Authors' comments:

[Pin Pin Tea-makorn](#) (PhD student at Stanford's Department of Electrical Engineering): "When we started this project, I was convinced that we would easily find evidence for the convergence in facial appearance. This is one of those theories that all undergrads learn in Psych 101. Also, it sounds rather intuitive: Spouses tend to spend much time together, and have similar hobbies and diets, so it is to be expected that they should grow more alike with time. Nevertheless, we were surprised that despite using a very large sample of facial images and very sensitive measurements of facial similarity, we could not find any evidence for convergence."

[Dr. Michal Kosinski](#) (Associate Professor of Organizational Behavior at Stanford's Graduate School of Business): "I think that scholars should spend more time checking each other's results and validating theories. Many are based on scant evidence and studies that would not be considered appropriate today. Much of what we are teaching to students, or advising to business leaders and policymakers, is based on weak evidence. Now, rejecting widely accepted hypotheses is surely not as exciting or as newsworthy as coming up with fancy new ones. Yet, it is as, if not more, important."

FAQ:

Q: What motivated you to undertake this research topic?

Pin Pin: In one of my first conversations with Michal, I made an observation that couples look alike. Michal, in turn, told me about this classical psychological theory claiming that couples grow more alike over time. We looked up the original study and were surprised that it was based on only 12 couples.

Q: What is the importance of this study?

Pin Pin: What predisposes two people to form and maintain a long-term romantic relationship is a fundamental question with critical consequences for the individuals involved, their families, and entire societies.

Q: Is part of the reason for the long-held belief that couples look more similar when they grow old together is that all older people look more similar than younger people, i.e., that aging reduces some of our faces' distinctiveness?

Pin Pin: That is likely. It seems that older people tend to look more similar to each other than younger people do. Thus, older couples likely look more similar to each other than younger couples do, but not because their facial appearance converges, but merely because older people look less distinct in general.

Q: Did you find evidence that people tend to form romantic relationships with similar-looking others?

Pin Pin: Yes, both our results and several previous studies showed that couples tend to be more similar to each other than random pairs of men and women. Thus far, it was believed that such similarity is driven by two factors: People's preference for similar others and spouses growing more alike with time. Our results show that the similarity is explained by the first factor.

Q: Why do people select partners who look similar to themselves?

Pin Pin: People often end up dating partners that are similar not only in their physical appearance but also in other traits, such as personality, values, and socioeconomic status. This observation is well documented in literature. Several explanations have been put forward. First, people tend to like others who are similar to themselves. Second, the rules of the dating market imply that people typically need to settle for a partner with a similar level of attractiveness (even if they would prefer to date someone more attractive). Finally, we are often surrounded by similar others, socially and geographically, so we naturally end up dating them.

Q: Can you give me some examples of couples whose photos were used in the study?

Pin Pin: Unfortunately, we cannot disclose the actual couples used in the study due to the privacy policy. We only have a sample stimulus set that we showed to our human judges (Figure 1).

Q: What is the next step in investigating this resemblance between couples?

Pin Pin: We suggest future studies done on different demographics such as homosexual and interracial couples. Another interesting line of research would be digging into which features of the couples' faces are similar.

On replication studies:

Replicability is almost universally accepted as the most important criterion of genuine scientific knowledge. (Rosenthal & Rosnow, 1984, p. 9)

Modern scientists are doing too much trusting and not enough verifying—to the detriment of the whole of science, and of humanity. ("How Science Goes Wrong," October 19, 2013, p. 13)

Pin Pin: I would like to highlight the value of revisiting old research questions and replicating published results. Even though our finding contradicts an older study, that doesn't mean that the previous study was wrong. In our case, the original study was very elegantly designed and one of the best of its time. However, as humans evolve, our psychological mechanisms definitely change. Therefore, I believe it is very important to update the results of studies from time to time, especially the ones that have not been widely replicated.

Michal: One of the major problems in social sciences is the pressure to come up with novel, amazing, newsworthy theories. This is how you get published, hired, and tenured. As a result, the field is filled with concepts and theories that are reclaimed, overhyped, or not validated properly. Consider, for example, the [hot and book-worthy "grit,"](#) which is, essentially, a [reappropriation of the well-known conscientiousness.](#) Or the [tens of well-known psychological theories that could not be replicated,](#) but that are still widely taught to students. Or leading social scientists openly admitting that they ["used data as a point of persuasion, and I never really worried about, 'Will this replicate or will this not?'"](#)

Importantly, most of the scholars act ethically and in good faith. Yet the incentives are set to encourage making novel (be it superficial) claims and discourage their validation. More than once, I have been asked: "Why would you rock the boat and try to ruin your colleague's career? Can't you come up with your own theory?" This is why I am so proud of Pin Pin for taking on this and other similar projects: Cleaning up the field might be the most important challenge faced by social scientists today, yet she is surely not going to get as many citations or as much recognition for her work as she would get if she came up with something new and flashy.

Figures

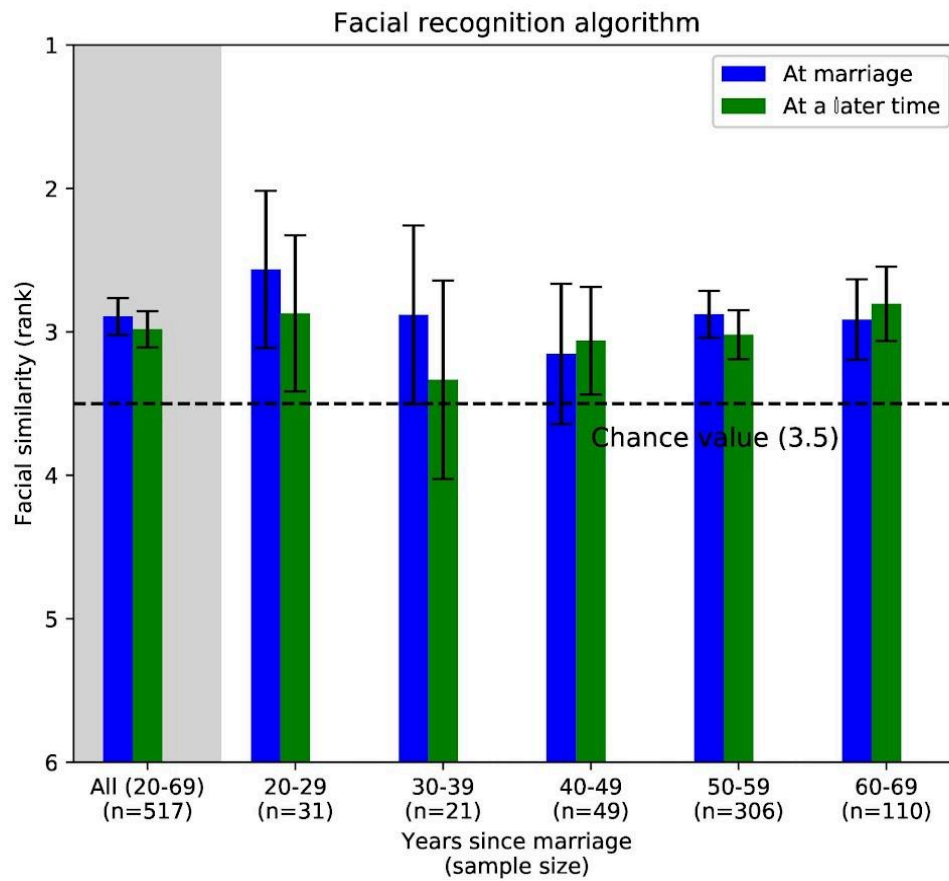


Figure 2: The average facial similarity of the spouses at marriage and 20 to 69 years later, estimated using the facial recognition algorithm. Error bars represent 95% confidence intervals.

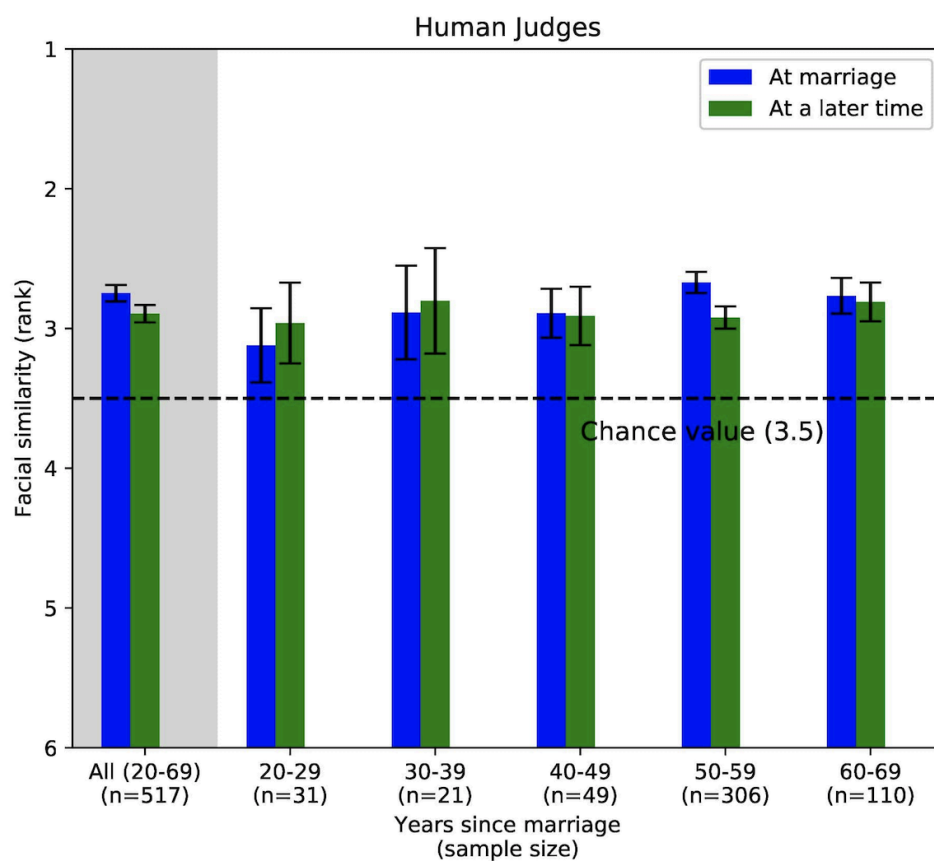


Figure 3: The average facial similarity of the spouses at marriage and 20 to 69 years later, estimated using human judgment. Error bars represent 95% confidence intervals.