

In press at the *Journal of Personality and Social Psychology*

Total length: 5 pages, double-spaced.

IN DEFENSE OF DILIGENCE:
A REJOINDER TO PELHAM AND CARVALLO.

Uri Simonsohn

The Wharton School

University of Pennsylvania

uws@wharton.upenn.edu

Abstract:

In Simonsohn (2011) I report the results from 14 studies that suggest all existing evidence of implicit egotism in marriage, job and location decisions is spurious. Lack of diligence by Pelham and colleagues explains in great part why the confounds behind their findings were not addressed in time. They almost never included controls, were dismissive of blatant alternative explanations, and on occasion misreported factual information that made confounds appear less important. Their rebuttal is similarly lacking in diligence. The specific empirical concerns it raises are contradicted by evidence, logic, or both. It reports misleading examples and inaccurate facts (some regarding the authors' own data). In this rejoinder I address all specific issues they raise about the empirics of my paper. I then provide perhaps the most striking example of lack of diligence in their earlier work. I close on a constructive note, providing two concrete suggestions on how to analyze data in future implicit egotism studies.

The rebuttal by Pelham and Carvallo (henceforth, PC) raises six specific objections to analyses in Simonsohn (2011). Here I show each of them is contradicted by evidence, logic or both. I then argue that the lack of diligence in the rebuttal is representative of their implicit egotism work more broadly. I justify this claim sharing, due to space constraints, only the perhaps most striking example (from the same initials marriage study in Jones, Pelham, Carvallo and Mirenberg, 2004). I close on a constructive note, providing two concrete suggestions on how to analyze data in future implicit egotism studies, avoiding errors from the past.

(1) Is “Does George stay in Georgia?” a valid test of implicit egotism? PC differentiate between short and long term consequences of name-similarity on the liking of a target, labeling them attraction and commitment respectively. They then argue that George *moving* to Georgia (what Pelham et al. (2002) tested) is about attraction, but George *staying* there (what I tested) is about commitment, and that implicit egotism makes strong predictions only about attraction.

The notion that implicit egotism is limited to attraction may be correct, but it cannot explain the incongruent findings. *All* of Pelham et al. field studies also focus only on commitment decisions. Their data classify George as a geoscientist if he has already published in that field, as moving to Georgia if he died there, and as liking Georgette if he married her.

(2) Would Maurine be a bad control for Josephine? Study 5 examined if people with similar first names (e.g., Joseph:Josephine) disproportionately married each other.¹ It used as controls, names that were highly correlated with target names in their spousal choices (e.g., Joseph and Earnest have similar distributions of spouse names so Earnest is one of Joseph’s controls. That is, Study 5 examined if Josephs marry Josephines more than Earnests do).

To avoid controls that had a name similarity to the target I did not use same initial ones (e.g., John could not be a control for Joseph). PC point out that I did not, however, exclude controls that were similar in other ways (e.g., Monica was a control for Erica). They then, perhaps disingenuously, argue that this may account for the disappearance of the name-similarity-effect in Study 5 since controls may also be

¹ In this paper, unless otherwise noted, study, figure and table numbers refer to Simonsohn (2011)

exhibiting implicit egotism. Below I discuss three problems with their argument (i) use of a misleading example, (ii) unfulfilled prediction, (iii) faulty logic.

(i) *Misleading example.* PC reprinted the control names only for Erica, who among the 24 target names is -by far- the most similar to its controls (Jessica, Monica and Vanessa). See full list in Table A4. Within that same name-pair, for instance, the control names for Eric are the rather different Christopher, Jonathan and Aaron. Even this cherry picked example, then, has high internal validity as Eric is more similar to Erica than Christopher, Jonathan and Aaron are.²

(ii) *Unfulfilled prediction.* Their argument predicts that name-pairs with very different controls should still exhibit the effect. They do not. Figure 3 shows, for example, that it is fully eliminated also for Joseph:Josephine with whom none of the six control names have the pronunciation of a single syllable in common (Paul, Patrick, Peter, Beatrice, Eva, Gloria).

(iii) *Faulty logic.* Name similarity in controls, to be a problem, must extend over to the opposite-sex target name. For example, imagine that the controls for Josephine had been Christine, Maurine and Nadine. While they all have the same ending, Joseph does not, so he should still egotistically predilect Josephine.

(3) *Is Guerra a bad control name for Guerrero?* In Study 3 I found that extremely similar Hispanic last names (e.g., Gonzalez:Gonzales) showed no evidence of implicit egotism. In an interview with the Wall Street Journal about the first draft of my paper (Bialik, 2010), Brett Pelham argued that this may be due to an aversion for names that one's is often misspelled as. In Simonsohn (2011) I responded to this hypothesis with three counterarguments. One of them was that we do observe implicit egotism among very similar *first* names e.g., Eric:Erica. PC countered that Eric would never be confused as Erica “unless [he] is fond of cross-dressing” (p.xx). My other two counterarguments, unaddressed by PC, were: “First, several last names in this study are *not* nearly identical (e.g., Mora:Morales or Aguilar:Aguirre) and we hence would expect a net effect of implicit egotism for them, but we do not see it. Second, it would be greatly coincidental that across different name pairs the negative effect of repulsion would always nearly

² Eric also is more similar to Erica than to Monica, Jessica and Vanessa, see point (iii).

exactly match the positive effect of implicit egotism, such that names with strong and weak same-last-name effects always have a very close to 0 very-similar-last-name effect” (Simonsohn 2011, pp.xxx).

(4) *Can New Yorkers freely choose where to live?* In Study 14 I found, using the entire voter registration file for New York (N=12 million), no evidence that people lived in addresses that resembled their birthdays. This finding is inconsistent with Pelham et al. (2002) where people disproportionately lived in towns with names resembling their birthday (N=495). PC propose this is in part because there are not enough available apartments in New York city (where 40% of the state population lives) for individuals to act upon their egotistical preferences. That is an empirical question. On Craigslist, on 01/05/2011 alone, there were more than 1000 listings for apartments with “3rd Avenue” in the title.³ That’s more than the entire birthday sample analyzed by Pelham et al. Furthermore, even if we assume that 40% of New Yorkers can’t easily choose where to live, the other 60% still can. The overall effect should be $(.4*0)+(.6*\text{something}) > 0$; it is not. It is =0.

(5) *Divorce and reverse causality.* In Study 4 I provided evidence of one of the multiple mechanisms hypothesized to be behind reverse causality in marrying someone with the same last name: remarrying one’s former spouse. PC claim (inaccurately) that their marriage data are all prior to 1920,⁴ and (also inaccurately) that divorce was “virtually nonexistent” in the 19th century.⁵ Despite these inexactitudes, the general point stands. If the only mechanism for reverse causality involved divorce, given that it was less common in the past, the concern is less extreme for some of their data. There are, of course, many other mechanisms. In my paper I mentioned: (i) marrying a relative, (ii) marrying abroad, immigrating, then remarrying in the US, (iii) marrying a relative of one’s dead husband, (iv) marrying legally after cohabitating for a long time, and (v) clerical errors. Most, probably *more* common in the 19th century.

(6) *Zeroing in vs. zeroing out Mr. and Mrs. Smith.* PC argue that in Study 1, where I teased apart marriages sharing an initial from those sharing an entire last name, I “overcorrected” for the latter by

³ Craigslist caps the number of hits at 1000.

⁴ Study 1b in Jones et al (2004) covers 1823-1965, about half the marriages in that sample occurred after 1920. A study described in pages 669-670 uses data from 2001.

⁵ The divorce rate for 1890 U.S. marriages was about 10% (Preston & McDonald, 1979).

employing a test “where the number of Smiths who married Smiths is set to [zero]” (p.xx). The test in question did no such thing. Being a standard independence χ^2 test on a contingency table, expected frequencies were obtained by multiplying the marginals. Here that implied assuming all last names, including Smith, are equally likely to marry any given last name, including Smith. In particular, I first created a table with last names of brides and grooms as rows and columns respectively, and where each cell contained the frequency of marriages between them. From those actual frequencies I then created the expected ones. Each table was then collapsed into two cells: one for observations sharing an initial (but not the full last name), and the other for everyone else (not sharing initials *or* sharing entire last name). A $\chi^2(1)$ was then performed on those two 2-cell tables.⁶

On the lack of diligence in previous research

Cohort, geography, ethnicity, religion and socioeconomic status, among others, are major drivers of both names people have and decisions they make. Diligently studying implicit egotism requires *simultaneously* addressing all these factors, plus the possibility of reverse causality.

Not *one* of the about 20 published field studies by Pelham and colleagues controls for more than one such factor, and most control for none. When they do, it is done qualitatively when it could be done quantitatively, often treating alternative explanations dismissively, and on at least two occasions providing erroneous yet easily available factual information that made confounds appear less likely than they truly were.

In the interest of brevity I share perhaps the most striking example of lack of diligence, encompassing all of the above in a single study. When examining if same-initial marriages are disproportionately common (Study 1 in Jones et al. (2004)) they address the ethnic confound concern by *arguing* that the source of the data, Walker county, being 94.6% White in the 1990 Census, was “ethnically homogenous.” First problem: the marriage data ends in 1920, 70 years before the race data were collected. In the 1890 Census that county was just 84% White. Second, they don’t disclose that the other county they analyze,

⁶ An alternative is to build a contingency table that excludes same last name marriages altogether (see footnote 11 in Simonsohn, 2011). That approach lead to results slightly *less* supportive of implicit egotism than those reported in Figure 1. This also does not assume 0% of Smiths marry a Smith; it makes no assumption on Smith-Smith.

Liberty (FL), was three times more diverse in 1990 (and just 56% White in 1890). Third, they unjustifiedly assume that 94.6% White is high enough to eliminate an ethnic confound. Study 3 in my paper shows that 99.89% is not. Fourth, there are no controls for other confounds such as: variation over time in initials (e.g., due to immigration), ethnicity beyond White/Black (e.g., at the time, Irish and Italian immigrants rarely inter-married), reverse causality, variation across towns in the county, religion, etc.

How to diligently study Implicit Egotism in the future?

Simonsohn (2011) does not falsify implicit egotism. It does invalidate the *existing* evidence of it in marriage, job, and location decisions. Future studies of implicit egotism in these or other settings must explicitly, quantitatively, systematically and simultaneously address the problems of reverse causality and likely name confounds. Two simple techniques described briefly below can do so for a large number of situations (see also supplemental materials in Simonsohn (*in press*)).

Reverse causality. One should avoid counting as evidence of implicit egotism identical name matches. Walt Disney worked for Walt Disney Productions and Hilary Clinton married Bill Clinton for reasons other than implicit egotism. Note however that focusing on similar rather than identical name matches *reduces* but not entirely eliminates reverse causality (e.g., Toyota was founded by Mr. Toyoda, and Louis is a popular baby name is Louisiana).

Confounds. To control for virtually *all* omitted variables at once, one can employ the approach I developed for studying the impact of first name similarity on marriage (Study 5): using as control names those most highly correlated with the target name in other instances of the dependent variable. For example, to examine if there are too many Thompsons marrying Johnsons one could use as controls last names most correlated with Thompson and most correlated with Johnson in the distribution of last names they marry.

Diligent data collection and analysis is what sets apart scientific from anecdotal evidence, let's embrace it unreservedly.

References

- Bialik, C. (2010, 15 May 2010). Are You W.S.J.? Is That Why You're Reading This?, *The Wall Street Journal*. Retrieved from <http://online.wsj.com/article/SB10001424052748704635204575242704020485696.html>
- Jones, J. T., Pelham, B. W., Carvalho, M., & Mirenberg, M. C. (2004). How Do I Love Thee? Let Me Count the Js: Implicit Egotism and Interpersonal Attraction. *Journal of Personality and Social Psychology*, 87(5), 665-683.
- Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie Sells Seashells by the Seashore: Implicit Egotism and Major Life Decisions. *Journal of Personality and Social Psychology*, 82(4), 469-487. doi: 10.1037//0022-3514.82.4.469
- Preston, S., & McDonald, J. (1979). The Incidence of Divorce within Cohorts of American Marriages Contracted since the Civil War. *Demography*, 16(1), 1-25.
- Simonsohn, U. (2011). Spurious? Name Similarity Effects (Implicit Egotism) in Marriage, Job, and Moving Decisions. *Journal of Personality and Social Psychology*.
- Simonsohn, U. (in press). Spurious Also? Name-Similarity-Effects (Implicit Egotism) in Employment Decisions. *Psychological Science*.