Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/3/11/e1700404/DC1

Supplementary Materials for

Authorship and contribution disclosures

Henry Sauermann and Carolin Haeussler

Published 15 November 2017, *Sci. Adv.* **3**, e1700404 (2017)
DOI: 10.1126/sciadv.1700404

This PDF file includes:

- table S16. Authorship positions and contributions controlling for quantity and quality of previous publications.
- table S17. Authorship positions and contributions using data from papers in the top 10% of article impact (citations).
- References (*36*–*38*)

Supplementary Materials

Supplementary Text

Conceptual background

When making decisions regarding which team members to include as co-authors and where to place them on the byline, scientists consider a range of factors (*6, 8*). We distinguish broadly between the "value" of team members' substantive contributions to the project, V_i, and "social" factors that are independent of actual contributions, S_i. The latter may include, for example, norms that senior members are listed as last authors regardless of their contributions, or teams' decision that first authorship goes to a junior team member who is on the job market rather than the member who made the most contributions (*6, 36*). Of course, factors such as seniority and hierarchical status may also determine which substantive contributions a particular team member makes (*2, 36*) and thus indirectly shape the value of the member's contributions to the project. Our conceptualization of "social" factors that may influence authorship includes only social mechanisms that are independent of actual contributions, while the value of contributions reflects all substantive contributions regardless of the underlying reasons.

Equation 1 describes an individual's author position on the byline, P_i, as a function of these two factors, V_i and S_i, as well as potential other influences (O_i).

$$P_i = f(V_i, S_i, O_i) \quad (1)$$

We are now interested in the degree to which an author's position on the byline can inform readers about the value of the author's substantive contributions. We examine this question empirically by studying how well author position, P_i, predicts different aspects of individuals' contributions, V_i. Before turning to the empirical analysis, however, it is useful to consider more explicitly different aspects of the value of contributions. This discussion will also be useful in understanding which aspects are observed in currently used contribution disclosures and which ones remain unobserved.

The prevailing approaches to interpreting authorship and disclosing author contributions conceptualize research as a production process that involves different activities such as designing an experiment, performing the experiment, analyzing data, and writing the paper (*2*). Each team member can be engaged in one or multiple of these activities, and can contribute a part or the entirety of a given activity. Figure S1 illustrates the resulting division of labor between N different team members, represented in the rows. Each column stands for one of K different activities required to complete the project. The variable a_{ij} indicates the share of activity j performed by member i, thus reflecting whether and at what level member i was engaged in activity j. We suggest that the value of a team member's contribution, V_i, depends on a_{ij} as well as on the importance of the different activities for the success of the overall project, I_j. A simple way to aggregate these aspects is a weighted average (see *37*):

$$V_i = \sum_{j=1}^{K} I_j \, a_{ij} \quad (2)$$

Equation 2 shows that the value of contributions is higher for individuals who are engaged in more activities, contribute a higher share of a given activity, or work on activities that are more important for project success.

Study 1: Additional Analyses

Controlling for quantity and quality of prior publications

Social factors such as the junior versus senior status of team members are likely to influence what types of activities team members perform (*19*) but may also influence how contribution statements are written for a given set of actual contributions (see our analysis of Study 2). The former is not a concern since contribution statements would still reflect actual contributions made. The latter, however, may introduce error and bias when using contribution statements as proxies of actual contributions. To partly address this issue, our regressions include a dummy variable indicating whether all authors have the same affiliation, which may be associated with stronger social influence of dominant team members. In addition, we now perform an auxiliary analysis that includes two factors that are likely to correlate positively with social status, namely measures of the quantity and quality of authors' prior publications. We obtained these measures from the Scopus database. We exclude from this analysis 940 papers that had at least one author who we could not match to the Scopus database; the records of all other authors were matched using the unique Scopus author identifier. We hand-checked authors with more than 200 returned publications and dropped some cases where the Scopus matching seemed erroneous. We also dropped a small number of papers with individuals for whom Scopus returned publications that were more than 60 years old. We use the log of the number of publications over the five years prior to the focal *PLOS ONE* article as a measure of the quantity of prior publications (i_lnpriorpubs_quantity) and the log of the average yearly number of citations to these articles (dividing total yearly citation counts by the total number of publications) as a proxy for the quality of prior publications (i_lnpriorpubs_quality). Tab. S1 shows that last authors have the highest scores on these measures, followed by middle and first authors.

Table S16 replicates our key regression models (table S2) with these additional controls included. While most of the results are qualitatively unchanged, we see two interesting differences: First authors are now more likely than last authors to have conceived, compared to a negative coefficient in the main models. Second, whereas first authors were estimated to be less likely to provide reagents/materials/analysis tools than last authors in the main models, this difference now disappears. Although these changes may reflect better controls for social influence (i.e., more accomplished authors may have pushed to be listed as having conceived and provided materials regardless of actual contributions), these changes are more likely to reflect the first mechanism discussed, i.e., that more accomplished scientists are indeed those who are more likely to perform these activities based on their experience and access to resources (*2, 24,*

38). Since we are not able to separate the impact of social status on the completion of contribution statements from its impact on actual division of labor, the regressions shown here likely control for "too much" and our main models are more suitable to answer our main research questions. Although not the focus of this analysis, we also briefly report the main effect of the new controls: The quantity of prior publications is positively associated with conceived, materials, wrote and other, but negatively associated with performed and analyzed. The quality of prior publications is positively associated with conceived, performed, analyzed, and wrote. Moreover, it is positively associated with the overall count of contributions.

Same contributions for all authors

We examine the special case that all authors are listed with the same contributions. A concern is that such contribution statements are inaccurate because authors state equal contributions simply to avoid difficult discussions and conflicts. We find that only 0.98% of papers state the same contributions for all authors. Moreover, this case is more common among small teams than large teams (e.g., 15.1% of papers with two authors, 3.3% of papers with three authors, and 1.7% of papers with four authors). These patterns give little reason for concern given that it is quite possible that all authors made the same types of contributions (though possibly with different shares of effort), and that this case would be expected especially in small teams. Excluding papers with all equal contributions does not change our substantive results (available upon request).

Top 10% of papers in terms of citations

PLOS ONE publishes a large number of papers, including many that have lower impact and may be of lower quality than papers published in more prestigious journals such as *Science*, *Nature*, or *PNAS*. To focus specifically on higher impact papers, we replicate our key regressions (Tab. S2) using only papers in the top 10% of annual citations (using article level metrics available on the *PLOS ONE* website as of Dec. 2015). The results are very similar to those obtained using the full sample (Tab. S17). One noticeable change is that the coefficient of "first author" becomes insignificant (but remains negative) in the regression of i_conceived&wrote (Model 8). Thus, the difference between first and last authors in terms of having performed these two activities appears to be somewhat smaller in high impact papers.

		Contribution type				
		c_1	.	.	.	c_k
Team member	m_1	a_{11}	.	.	.	a_{1k}
	.	.	.	.	.	.
	.	.	.	.	.	.
	.	.	.	.	.	.
	.	.	.	.	.	.
	m_n	a_{n1}	.	.	.	a_{nk}

fig. S1. Team members and their respective contributions (schematic). Each column stands for one of K different activities required to complete the project. The variable a_{ij} indicates the share of activity j performed by member i, thus reflecting whether and at what level member i was engaged in activity j.

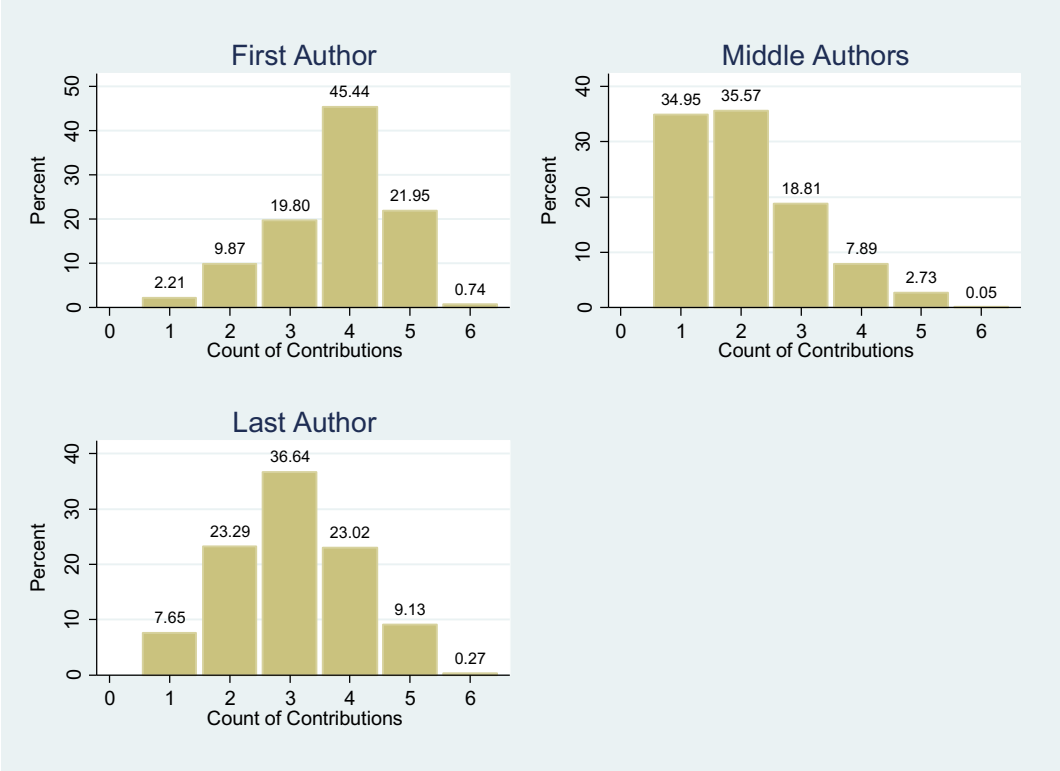

fig. S2. Distribution of the count of contributions by position (teams of six).

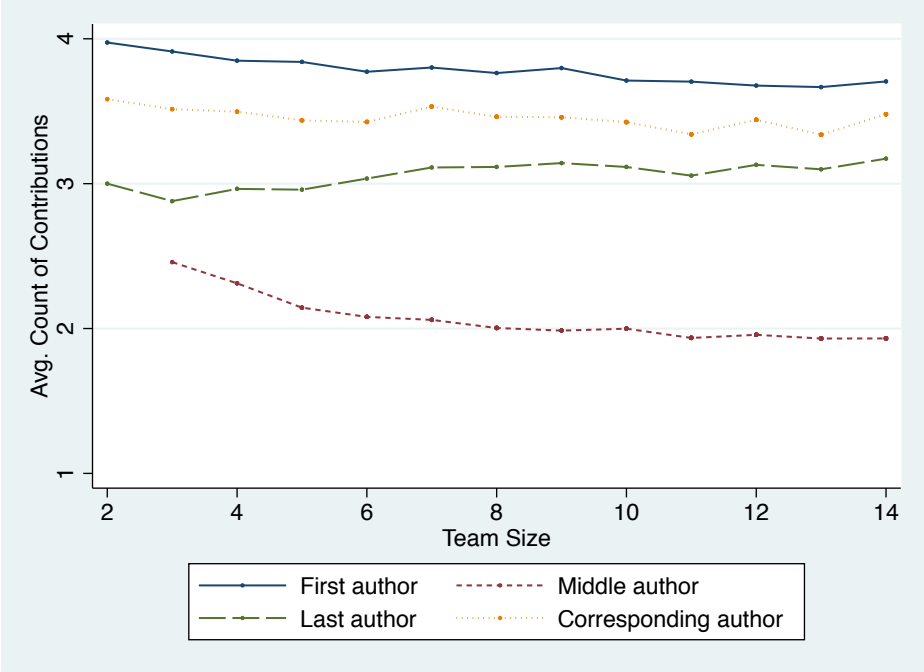

fig. S3. Count of contributions by team size.

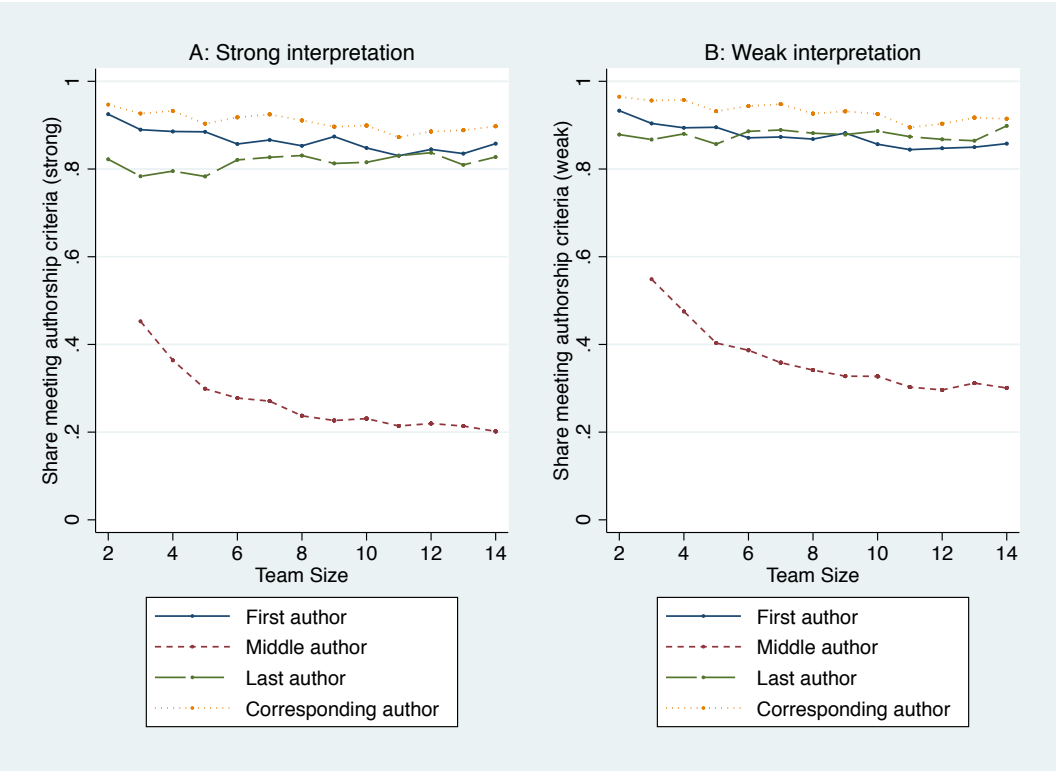

fig. S4. Share of authors who fulfill ICMJE authorship criteria. Using the strong interpretation of criteria (A) and using weak interpretation (B). By team size.

table S1. Summary statistics for Study 1.

	Full sample				First authors	Middle authors	Last authors	Corresponding authors
	Mean	SD	Min	Max	Mean	Mean	Mean	Mean
Individual level (N=79,776)								
i_conceived	0.49		0	1	0.80	0.32	0.87	0.93
i_performed	0.51		0	1	0.86	0.50	0.23	0.40
i_materials	0.34		0	1	0.28	0.35	0.35	0.35
i_analyzed	0.59		0	1	0.94	0.49	0.65	0.80
i_wrote	0.48		0	1	0.88	0.29	0.84	0.93
i_other	0.09		0	1	0.05	0.09	0.09	0.07
i_countcontributions	2.49	1.25	1	6	3.82	2.04	3.03	3.47
First author	0.16		0	1	1.00	0.00	0.00	0.32
Middle author	0.68		0	1	0.00	1.00	0.00	0.09
Last author	0.16		0	1	0.00	0.00	1.00	0.59
Corresponding author	0.18		0	1	0.35	0.02	0.65	1.00
i_lnpriorpubs_quantity*	2.02	1.27	0	5.30	1.49	1.95	2.81	2.50
i_lnpriorpubs_quality*	1.39	0.82	0	6.72	1.27	1.36	1.62	1.58
i_icmjefulfilled_strong	0.44		0	1	0.87	0.26	0.81	0.91
i_icmjefulfilled_weak	0.52		0	1	0.88	0.36	0.88	0.94
Article level (N=12,772)								
t_teamsize	6.25	3.02	2	14				
t_totalactivitieslisted	4.80	0.67	1	6				
t_alphaorder	0.07		0	1				
t_affiliations_d	0.75		0	1				
t_biochemistry	0.14		0	1				
t_biophysics	0.04		0	1				
t_biotechnology	0.05		0	1				
t_cardiovasculars	0.04		0	1				
t_cell_Biology	0.24		0	1				
t_chemical_Biology	0.03		0	1				
t_computational_Biology	0.09		0	1				
t_developmental_Biology	0.08		0	1				
t_diabetes_and_Endocrinology	0.04		0	1				
t_ecology	0.09		0	1				
t_evidence_Based_Healthcare	0.01		0	1				
t_evolutionary_Biology	0.10		0	1				
t_gastroenterology	0.02		0	1				
t_genetics_and_Genomics	0.21		0	1				
t_hematology	0.02		0	1				
t_immunology	0.12		0	1				
t_infectious_Diseases	0.16		0	1				
t_marine_and_Aquatic_Sciences	0.02		0	1				
t_mental_Health	0.02		0	1				
t_microbiology	0.10		0	1				
t_molecular_Biology	0.12		0	1				
t_neurological_Disorders	0.06		0	1				
t_neuroscience	0.15		0	1				
t_non_Clinical_Medicine	0.01		0	1				
t_nutrition	0.02		0	1				
t_oncology	0.08		0	1				
t_pathology	0.03		0	1				
t_pediatrics_and_Child_Health	0.02		0	1				
t_pharmacology	0.03		0	1				
t_physiology	0.08		0	1				
t_plant_Biology	0.02		0	1				
t_public_Health	0.08		0	1				
t_virology	0.06		0	1				
t_otherbiolife	0.07		0	1				

Note: * is based on 73,847 individual author observations that could be matched to Scopus data

table S2. Authorship positions and contributions.

	Full sample							Full sample			Non-alphabetical	Alphabetical
	1	2	3	4	5	6	7	8	9	10	11	12
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	Poisson
	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_conceived& wrote	i_conceived& wrote& analzyed	i_performed & analyzed	i_count contributions	i_count contributions
VARIABLES												
First author	0.232**	-0.065**	0.637**	-0.064**	0.289**	0.038**	-0.044**	-0.018**	0.171**	0.632**	0.275**	0.283**
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.006)	(0.006)	(0.005)	(0.006)	(0.013)
Middle author	-0.372**	-0.531**	0.273**	-0.008	-0.152**	-0.516**	0.003	-0.581**	-0.433**	0.086**	-0.261**	-0.204**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.002)	(0.004)	(0.005)	(0.004)	(0.007)	(0.025)
Last author	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_teamsize	-0.034**	-0.018**	0.004	-0.032**	-0.014**	-0.027**	-0.011**	-0.022**	-0.015**	-0.016**	0.023	-0.194**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.024)	(0.059)
t_teamsize_sq	0.001**	0.000*	0.000	0.001**	0.001	0.001**	0.000	0.001**	0.001**	0.001**	-0.009**	0.025**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.010)
t_totalactivitieslisted	0.108**	-0.000	0.008	0.216**	-0.022**	-0.045**	0.113**	-0.010**	0.003	0.022**	0.132**	0.172**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)	(0.005)	(0.010)
t_alphaorder	-0.010	-0.014	0.005	0.009	-0.009	-0.023*	0.011	-0.007	-0.006	0.001		
	(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)	(0.009)	(0.010)	(0.011)	(0.010)		
t_affiliations_d	0.013*	0.025**	-0.089**	0.054**	-0.001	0.046**	-0.002	0.025**	0.010	-0.041**	0.000	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.016)
t_published	0.000**	0.000**	0.000	0.000**	0.000**	0.000*	-0.000**	0.000**	0.000**	0.000**	0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	0.273*	0.597**	0.102	-0.997**	0.441**	0.878**	-0.029	0.588**	0.336**	-0.170	0.008	0.192
	(0.120)	(0.112)	(0.113)	(0.130)	(0.128)	(0.119)	(0.111)	(0.096)	(0.086)	(0.109)	(0.164)	(0.344)
Observations	79,776	79,776	79,776	79,776	79,776	79,776	79,776	79,776	79,776	79,776	19,519	2,217
R-squared		0.240	0.154	0.086	0.123	0.304	0.092	0.339	0.316	0.219		

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. Models 11-12 use only individuals in teams with <6 members. LPM, linear probability model (OLS).

table S3. Authorship and contributions for teams with 2, 6, and 10 authors.

	Teamsize = 2							Teamsize = 6							Teamsize = 10						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	LPM	LPM	LPM	LPM	LPM	LPM
	i_count	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_count	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_count	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other
VARIABLES	contributions							contributions							contributions						
First author	0.281**	0.036*	0.608**	-0.025	0.320**	0.068**	-0.032**	0.218**	-0.103**	0.637**	-0.071**	0.304**	0.013	-0.043**	0.175**	-0.084**	0.601**	-0.089**	0.217**	0.017	-0.065**
	(0.011)	(0.016)	(0.017)	(0.015)	(0.016)	(0.013)	(0.007)	(0.011)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.007)	(0.016)	(0.020)	(0.020)	(0.021)	(0.019)	(0.020)	(0.011)
Middle author								-0.378**	-0.543**	0.259**	-0.006	-0.143**	-0.529**	0.007	-0.444**	-0.573**	0.267**	-0.007	-0.222**	-0.569**	-0.013
								(0.011)	(0.010)	(0.013)	(0.012)	(0.013)	(0.011)	(0.006)	(0.015)	(0.014)	(0.018)	(0.018)	(0.018)	(0.016)	(0.010)
Last author	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_totalactivitieslisted	0.211**	0.078**	0.159**	0.384**	0.052**	-0.013	0.037**	0.097**	0.010	-0.005	0.212**	-0.026*	-0.062**	0.111**	0.062**	-0.033	-0.030	0.162**	-0.043*	-0.058**	0.143**
	(0.007)	(0.015)	(0.011)	(0.017)	(0.014)	(0.008)	(0.014)	(0.009)	(0.011)	(0.011)	(0.012)	(0.012)	(0.010)	(0.013)	(0.018)	(0.019)	(0.017)	(0.018)	(0.019)	(0.017)	(0.021)
t_alphaorder	-0.006	0.007	0.004	-0.033	0.009	-0.016	0.009	-0.194	-0.171	-0.172	-0.232**	-0.039	-0.188	0.361	0.422*	0.618**	0.547**	0.275*	0.084	-0.218*	-0.012
	(0.010)	(0.015)	(0.015)	(0.017)	(0.015)	(0.012)	(0.011)	(0.192)	(0.131)	(0.228)	(0.055)	(0.197)	(0.175)	(0.302)	(0.171)	(0.120)	(0.130)	(0.133)	(0.181)	(0.092)	(0.044)
t_affiliations_d	0.005	0.001	-0.017	0.044	-0.014	-0.000	0.012	0.006	0.027	-0.133**	0.067**	-0.012	0.063**	0.005	0.176**	0.103**	-0.117**	0.121**	0.175**	0.115**	-0.038
	(0.013)	(0.019)	(0.020)	(0.023)	(0.019)	(0.015)	(0.015)	(0.018)	(0.017)	(0.016)	(0.017)	(0.020)	(0.017)	(0.013)	(0.049)	(0.031)	(0.037)	(0.044)	(0.040)	(0.035)	(0.034)
t_published	0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000*	0.000	0.000	0.000	-0.000**	0.000**	0.000	0.000	0.000*	0.000*	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field Fixed Effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	-0.507*	-0.158	-0.842*	-1.654**	-0.077	0.650*	-0.246	0.305	0.607*	-0.238	-1.087**	0.616	0.778*	0.397	-1.207*	0.252	0.102	-1.841**	-0.320	0.025	-0.285
	(0.252)	(0.372)	(0.345)	(0.427)	(0.382)	(0.292)	(0.303)	(0.299)	(0.308)	(0.297)	(0.333)	(0.336)	(0.306)	(0.307)	(0.507)	(0.453)	(0.427)	(0.549)	(0.498)	(0.466)	(0.401)
Observations	2,238	2,238	2,238	2,238	2,238	2,238	2,238	8,940	8,940	8,940	8,940	8,940	8,940	8,940	7,040	7,040	7,040	7,040	7,040	7,040	7,040
R-squared		0.048	0.447	0.370	0.212	0.041	0.053		0.227	0.175	0.109	0.133	0.279	0.100		0.202	0.107	0.062	0.110	0.264	0.136

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. LPM, linear probability model (OLS).

table S4. Incidence of alphabetical authorship.

Team size	Alphabetical order predicted by chance	Alphabetical order observed in the data		
	Share	Share	95% Conf. interval	
2	50.000000000%	50.58%	47.65%	53.51%
3	16.666666667%	15.60%	13.79%	17.41%
4	4.166666667%	4.04%	3.09%	4.98%
5	0.833333333%	1.16%	0.64%	1.68%
6	0.138888889%	0.13%	-0.05%	0.32%
7	0.019841270%	0.15%	-0.06%	0.36%
8	0.002480159%	0.00%		
9	0.000275573%	0.00%		
10	0.000027557%	0.14%	-0.14%	0.42%
11	0.000002505%	0.00%		
12	0.000000209%	0.00%		
13	0.000000016%	0.00%		
14	0.000000001%	0.00%		

Note: Alphabetical order predicted by chance is computed based on the number of permutations of n distinct names, where n equals team size.

table S5. Authorship position and corresponding author status.

	Non-corresponding	Corresponding	Total
First author	8,331	4,441	12,772
	65%	35%	100%
	13%	32%	16%
Middle author	53,000	1,232	54,232
	98%	2%	100%
	81%	9%	68%
Last author	4,480	8,292	12,772
	35%	65%	100%
	7%	59%	16%
Total	65,811	13,965	79,776
	82%	18%	100%
	100%	100%	100%

table S6. Corresponding author status and contributions.

	Full sample							First	Middle	Last
	1	2	3	4	5	6	7	8	9	10
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	Poisson	Poisson
VARIABLES	i_countcontributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_countcontributions	i_countcontributions	i_countcontributions
Corresponding author	0.170**	0.255**	-0.116**	0.051**	0.139**	0.246**	-0.020**	0.097**	0.380**	0.196**
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)	(0.011)	(0.008)
First author	0.282**	0.012*	0.602**	-0.048**	0.331**	0.112**	-0.050**			
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)			
Middle author	-0.265**	-0.372**	0.201**	0.024**	-0.065**	-0.363**	-0.010**			
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)			
Last author	omitted	omitted	omitted	omitted	omitted	omitted	omitted			
t_teamsize	-0.032**	-0.016**	0.003	-0.031**	-0.013**	-0.026**	-0.011**	-0.026**	-0.075**	-0.022**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.007)	(0.005)
t_teamsize_sq	0.001**	0.000	0.000	0.001**	0.000	0.001**	0.000	0.001**	0.003**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
t_totalactivitieslisted	0.108**	-0.000	0.008	0.216**	-0.023**	-0.045**	0.113**	0.136**	0.087**	0.122**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.003)	(0.006)	(0.005)
t_alphaorder	-0.011	-0.015	0.005	0.009	-0.010	-0.024*	0.011	-0.001	0.030	0.024
	(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)	(0.009)	(0.008)	(0.026)	(0.013)
t_affiliations_d	0.010	0.024**	-0.089**	0.054**	-0.001	0.046**	-0.002	-0.013*	0.024*	0.021**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.010)	(0.008)
t_published	0.000**	0.000**	0.000	0.000**	0.000**	0.000*	-0.000**	0.000*	0.000**	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	0.155	0.429**	0.178	-1.031**	0.348**	0.715**	-0.016	0.556**	-0.004	0.071
	(0.120)	(0.113)	(0.113)	(0.130)	(0.128)	(0.119)	(0.111)	(0.105)	(0.179)	(0.151)
Observations	79,776	79,776	79,776	79,776	79,776	79,776	79,776	12,772	54,232	12,772
R-squared		0.263	0.159	0.087	0.130	0.325	0.093			

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. LPM, linear probability model (OLS).

table S7. Predicted likelihood of particular contributions and predicted contribution counts.

	First authors	Middle authors	Last authors
Individual level (N=79,776)			
i_conceived	0.79	0.33	0.86
i_performed	0.86	0.50	0.23
i_materials	0.29	0.35	0.35
i_analyzed	0.93	0.49	0.64
i_wrote	0.86	0.30	0.82
i_other	0.05	0.09	0.09
i_countcontributions	3.77	2.06	2.99

Note: Calculated based on the regression models 1-7 reported in Tab. S2.

table S8. Author contributions by position and team size.

	First authors							Middle authors						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	LPM	LPM	LPM	LPM	LPM	LPM
VARIABLES	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other
t_teamsize=2	0.107**	0.113**	0.048**	0.137**	0.026*	0.099**	-0.014							
	(0.010)	(0.017)	(0.014)	(0.019)	(0.011)	(0.014)	(0.010)							
t_teamsize=3	0.065**	0.075**	0.036**	0.086**	0.017*	0.041**	-0.006	0.200**	0.152**	-0.040*	0.061**	0.075**	0.185**	0.013
	(0.009)	(0.015)	(0.012)	(0.016)	(0.009)	(0.012)	(0.008)	(0.015)	(0.016)	(0.016)	(0.014)	(0.016)	(0.016)	(0.010)
t_teamsize=4	0.035**	0.034*	0.002	0.059**	0.003	0.033**	0.002	0.124**	0.068**	-0.021	0.042**	0.062**	0.099**	0.020*
	(0.009)	(0.014)	(0.011)	(0.015)	(0.009)	(0.012)	(0.008)	(0.013)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.009)
t_teamsize=5	0.016	0.034*	-0.015	0.016	-0.001	0.029*	-0.002	0.032*	0.017	-0.013	0.007	0.022	0.028*	0.005
	(0.009)	(0.014)	(0.011)	(0.015)	(0.008)	(0.012)	(0.008)	(0.013)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.009)
t_teamsize=6	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_teamsize=7	-0.002	-0.004	-0.008	0.009	-0.006	0.007	-0.010	-0.018	-0.015	0.011	0.005	-0.005	-0.005	-0.029**
	(0.010)	(0.016)	(0.012)	(0.016)	(0.009)	(0.013)	(0.008)	(0.013)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.009)
t_teamsize=8	-0.024*	-0.019	-0.020	-0.025	-0.020*	-0.009	-0.001	-0.056**	-0.022	0.006	-0.031**	-0.009	-0.040**	-0.022*
	(0.010)	(0.017)	(0.013)	(0.018)	(0.010)	(0.014)	(0.009)	(0.014)	(0.011)	(0.011)	(0.011)	(0.013)	(0.012)	(0.009)
t_teamsize=9	-0.012	-0.015	-0.024	-0.018	-0.008	0.012	0.005	-0.064**	-0.050**	0.030**	-0.016	-0.018	-0.054**	-0.023*
	(0.011)	(0.018)	(0.013)	(0.019)	(0.010)	(0.014)	(0.010)	(0.014)	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.010)
t_teamsize=10	-0.031**	0.010	-0.049**	-0.024	-0.034**	-0.007	-0.017	-0.060**	-0.031*	0.031*	-0.024	-0.024	-0.046**	-0.029**
	(0.012)	(0.018)	(0.016)	(0.020)	(0.012)	(0.016)	(0.009)	(0.016)	(0.013)	(0.012)	(0.013)	(0.014)	(0.013)	(0.010)
t_teamsize=11	-0.038**	-0.016	-0.043*	-0.024	-0.026	-0.033	-0.005	-0.088**	-0.052**	0.027*	-0.026	-0.028	-0.069**	-0.030**
	(0.014)	(0.021)	(0.018)	(0.023)	(0.013)	(0.019)	(0.011)	(0.017)	(0.013)	(0.014)	(0.014)	(0.015)	(0.014)	(0.011)
t_teamsize=12	-0.051**	-0.005	-0.064**	-0.037	-0.046**	-0.021	-0.021	-0.088**	-0.062**	0.055**	-0.026	-0.021	-0.066**	-0.058**
	(0.015)	(0.023)	(0.020)	(0.025)	(0.016)	(0.020)	(0.011)	(0.019)	(0.014)	(0.015)	(0.016)	(0.017)	(0.016)	(0.011)
t_teamsize=13	-0.044*	-0.021	-0.029	-0.065*	-0.037*	-0.030	0.011	-0.097**	-0.073**	0.048**	-0.038*	-0.017	-0.082**	-0.034*
	(0.018)	(0.028)	(0.022)	(0.028)	(0.018)	(0.024)	(0.016)	(0.022)	(0.015)	(0.016)	(0.018)	(0.020)	(0.017)	(0.015)
t_teamsize=14	-0.045*	-0.008	-0.070*	-0.070*	-0.033	-0.007	0.017	-0.097**	-0.064**	0.052**	-0.041*	-0.021	-0.082**	-0.040**
	(0.021)	(0.031)	(0.029)	(0.033)	(0.021)	(0.026)	(0.019)	(0.023)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.015)
t_totalactivitieslisted	0.139**	0.067**	0.128**	0.230**	0.043**	0.019**	0.034**	0.088**	-0.023**	-0.031**	0.200**	-0.044**	-0.054**	0.132**
	(0.003)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.007)
t_alphaorder	-0.009	-0.013	-0.011	-0.014	0.005	-0.022	0.018*	0.008	0.011	0.006	-0.029	0.012	-0.001	0.017
	(0.008)	(0.015)	(0.012)	(0.016)	(0.009)	(0.012)	(0.009)	(0.026)	(0.027)	(0.026)	(0.025)	(0.026)	(0.025)	(0.018)
t_affiliations_d	0.000	0.007	-0.044**	0.015	-0.003	0.030**	-0.001	0.027**	0.051**	-0.147**	0.081**	0.000	0.073**	-0.001
	(0.006)	(0.009)	(0.007)	(0.010)	(0.006)	(0.008)	(0.005)	(0.010)	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)	(0.006)
t_published	0.000	-0.000	0.000	0.000	0.000	0.000**	-0.000**	0.000**	0.000**	0.000	0.000	0.000**	0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	0.466**	0.592**	0.028	-1.200**	0.594**	0.272	0.283**	-0.344	-0.052	0.661**	-0.989**	0.206	0.231	-0.169
	(0.108)	(0.179)	(0.148)	(0.191)	(0.109)	(0.145)	(0.106)	(0.178)	(0.144)	(0.147)	(0.159)	(0.166)	(0.156)	(0.134)
Observations	12,772	12,772	12,772	12,772	12,772	12,772	12,772	54,232	54,232	54,232	54,232	54,232	54,232	54,232
R-squared		0.024	0.151	0.118	0.022	0.040	0.036		0.021	0.032	0.071	0.015	0.058	0.105

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. LPM, linear probability model (OLS).

table S9. Types of contributions by position and team size.

	Last authors							Corresponding						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM	Poisson	LPM	LPM	LPM	LPM	LPM	LPM
VARIABLES	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other
t_teamsize=2	0.072**	-0.039*	0.055**	0.123**	0.052*	0.010	0.015	0.123**	0.020	0.141**	0.143**	0.093**	0.030**	0.002
	(0.016)	(0.018)	(0.020)	(0.020)	(0.023)	(0.017)	(0.013)	(0.013)	(0.012)	(0.021)	(0.019)	(0.019)	(0.011)	(0.011)
t_teamsize=3	0.002	-0.052**	-0.005	0.049**	0.012	-0.022	0.026*	0.067**	0.016	0.079**	0.078**	0.046**	0.008	0.004
	(0.013)	(0.014)	(0.016)	(0.016)	(0.018)	(0.014)	(0.011)	(0.010)	(0.010)	(0.017)	(0.016)	(0.015)	(0.009)	(0.009)
t_teamsize=4	0.007	-0.021	-0.007	0.025	0.023	-0.013	0.016	0.043**	0.003	0.032	0.053**	0.041**	0.016	0.006
	(0.012)	(0.012)	(0.015)	(0.016)	(0.017)	(0.013)	(0.010)	(0.010)	(0.009)	(0.017)	(0.016)	(0.014)	(0.008)	(0.009)
t_teamsize=5	-0.020	-0.028*	-0.032*	0.024	-0.006	-0.025	0.009	0.004	-0.005	-0.005	0.033*	0.004	-0.014	0.000
	(0.012)	(0.012)	(0.015)	(0.016)	(0.017)	(0.013)	(0.010)	(0.010)	(0.009)	(0.016)	(0.016)	(0.015)	(0.009)	(0.009)
t_teamsize=6	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_teamsize=7	0.008	-0.010	0.002	0.023	0.022	0.009	-0.021*	0.017	-0.004	0.012	0.028	0.025	0.003	-0.008
	(0.013)	(0.012)	(0.016)	(0.017)	(0.017)	(0.013)	(0.010)	(0.011)	(0.009)	(0.017)	(0.017)	(0.015)	(0.009)	(0.009)
t_teamsize=8	-0.004	0.023	-0.016	-0.039*	0.031	0.011	-0.021	-0.018	-0.003	-0.013	-0.043*	0.019	-0.009	-0.013
	(0.014)	(0.012)	(0.017)	(0.019)	(0.019)	(0.014)	(0.011)	(0.012)	(0.010)	(0.018)	(0.018)	(0.016)	(0.011)	(0.010)
t_teamsize=9	0.003	0.023	-0.011	-0.005	0.020	-0.009	-0.006	-0.018	-0.006	-0.027	-0.024	0.007	-0.024*	0.009
	(0.015)	(0.012)	(0.018)	(0.020)	(0.020)	(0.016)	(0.012)	(0.013)	(0.010)	(0.020)	(0.019)	(0.017)	(0.012)	(0.011)
t_teamsize=10	-0.007	-0.013	0.002	-0.027	0.027	-0.005	-0.009	-0.024	-0.010	-0.027	-0.018	0.002	-0.019	-0.016
	(0.016)	(0.014)	(0.019)	(0.021)	(0.021)	(0.017)	(0.013)	(0.014)	(0.012)	(0.021)	(0.021)	(0.018)	(0.013)	(0.010)
t_teamsize=11	-0.027	0.018	-0.031	-0.082**	0.036	0.006	-0.032*	-0.055**	-0.023	-0.050*	-0.066**	0.007	-0.044**	-0.016
	(0.018)	(0.015)	(0.021)	(0.023)	(0.024)	(0.019)	(0.014)	(0.016)	(0.013)	(0.023)	(0.023)	(0.020)	(0.015)	(0.013)
t_teamsize=12	-0.011	0.015	-0.006	-0.022	0.017	0.024	-0.061**	-0.035*	0.001	-0.028	-0.019	-0.020	-0.037*	-0.017
	(0.019)	(0.017)	(0.024)	(0.027)	(0.027)	(0.020)	(0.014)	(0.018)	(0.013)	(0.027)	(0.027)	(0.023)	(0.017)	(0.016)
t_teamsize=13	-0.010	0.017	-0.013	-0.023	0.010	-0.005	-0.017	-0.058**	-0.007	-0.080**	-0.061*	-0.038	-0.029	0.009
	(0.023)	(0.019)	(0.027)	(0.030)	(0.031)	(0.025)	(0.019)	(0.019)	(0.016)	(0.028)	(0.029)	(0.028)	(0.018)	(0.018)
t_teamsize=14	0.004	0.002	-0.018	-0.047	0.067	0.024	-0.014	-0.027	-0.005	-0.038	-0.062	0.013	-0.021	0.018
	(0.026)	(0.024)	(0.032)	(0.035)	(0.035)	(0.027)	(0.022)	(0.023)	(0.018)	(0.035)	(0.035)	(0.029)	(0.022)	(0.021)
t_totalactivitieslisted	0.119**	0.011	0.037**	0.264**	-0.007	-0.078**	0.121**	0.147**	0.045**	0.110**	0.272**	0.024**	-0.030**	0.078**
	(0.005)	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.006)
t_alphaorder	-0.002	-0.006	0.012	-0.012	-0.005	-0.012	0.018	-0.001	-0.001	0.008	-0.033*	0.001	0.001	0.017
	(0.014)	(0.016)	(0.018)	(0.016)	(0.020)	(0.015)	(0.011)	(0.011)	(0.011)	(0.018)	(0.016)	(0.016)	(0.009)	(0.010)
t_affiliations_d	0.004	-0.020*	0.019	0.021*	-0.003	-0.001	-0.002	0.036**	-0.006	0.063**	0.023*	0.033**	0.017**	-0.005
	(0.008)	(0.008)	(0.010)	(0.010)	(0.011)	(0.009)	(0.007)	(0.007)	(0.006)	(0.011)	(0.010)	(0.010)	(0.006)	(0.006)
t_published	0.000*	0.000*	0.000	0.000**	0.000	0.000	-0.000*	-0.000	-0.000	0.000	0.000**	0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	0.118	0.440**	-0.028	-1.777**	0.507*	1.126**	-0.157	0.475**	0.712**	-0.246	-1.522**	0.457*	1.249**	0.240
	(0.157)	(0.154)	(0.186)	(0.198)	(0.213)	(0.167)	(0.135)	(0.129)	(0.117)	(0.205)	(0.197)	(0.178)	(0.117)	(0.123)
Observations	12,772	12,772	12,772	12,772	12,772	12,772	12,772	13,965	13,965	13,965	13,965	13,965	13,965	13,965
R-squared		0.045	0.018	0.142	0.064	0.038	0.098		0.025	0.067	0.142	0.024	0.019	0.06

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. LPM, linear probability model (OLS).

table S10. ICMJE authorship criteria fulfilled by position and team size.

	Full sample			
	1	2	3	4
	LPM	LPM	LPM	LPM
VARIABLES	i_icmjefulfilled_strong	i_icmjefulfilled_strong	i_icmjefulfilled_weak	i_icmjefulfilled_weak
Corresponding		0.266**		0.214**
		(0.005)		(0.004)
First author	0.064**	0.144**	0.007	0.072**
	(0.005)	(0.005)	(0.004)	(0.004)
Middle author	-0.527**	-0.362**	-0.494**	-0.361**
	(0.004)	(0.005)	(0.004)	(0.005)
Last author	omitted	omitted	omitted	omitted
t_teamsize=2	0.047**	0.039**	0.063**	0.057**
	(0.012)	(0.012)	(0.011)	(0.011)
t_teamsize=3	0.060**	0.057**	0.076**	0.073**
	(0.010)	(0.010)	(0.010)	(0.010)
t_teamsize=4	0.046**	0.045**	0.061**	0.060**
	(0.009)	(0.009)	(0.009)	(0.009)
t_teamsize=5	0.012	0.011	0.011	0.010
	(0.008)	(0.008)	(0.009)	(0.009)
t_teamsize=6	omitted	omitted	omitted	omitted
t_teamsize=7	-0.000	0.000	-0.022*	-0.021*
	(0.009)	(0.009)	(0.009)	(0.009)
t_teamsize=8	-0.027**	-0.026**	-0.047**	-0.046**
	(0.009)	(0.009)	(0.010)	(0.010)
t_teamsize=9	-0.037**	-0.036**	-0.059**	-0.058**
	(0.009)	(0.009)	(0.010)	(0.010)
t_teamsize=10	-0.038**	-0.036**	-0.059**	-0.057**
	(0.010)	(0.010)	(0.011)	(0.011)
t_teamsize=11	-0.050**	-0.048**	-0.083**	-0.082**
	(0.011)	(0.011)	(0.012)	(0.012)
t_teamsize=12	-0.046**	-0.044**	-0.097**	-0.095**
	(0.013)	(0.013)	(0.014)	(0.014)
t_teamsize=13	-0.059**	-0.057**	-0.089**	-0.087**
	(0.014)	(0.014)	(0.017)	(0.017)
t_teamsize=14	-0.060**	-0.058**	-0.095**	-0.094**
	(0.016)	(0.016)	(0.019)	(0.019)
t_totalactivitieslisted	-0.029**	-0.030**	0.057**	0.057**
	(0.004)	(0.004)	(0.005)	(0.005)
t_alphaorder	-0.006	-0.006	0.008	0.008
	(0.011)	(0.011)	(0.010)	(0.010)
t_affiliations_d	0.031**	0.030**	0.037**	0.036**
	(0.006)	(0.006)	(0.006)	(0.006)
t_published	0.000**	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl
Constant	0.620**	0.450**	0.467**	0.331**
	(0.112)	(0.112)	(0.126)	(0.127)
Observations	79,776	79,776	79,776	79,776
R-squared	0.318	0.343	0.278	0.294

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. Models 1-2 use strict definition, models 3-4 use weaker definition (more permissive). LPM, linear probability model (OLS).

table S11. Summary statistics for Study 2.

	Full sample (N=6,002)					Bio/Life Sciences (N=3,872)	Medical/Health Sciences (N=708)	Physical Sciences (N=668)	Social Sciences (N=257)	Other Fields (N=497)	Junior Scientists (N=1,573)	Senior Scientists (N=3,729)	PLOS ONE (N=3,800)	PNAS (2,202)
	N	Mean	SD	Min	Max	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Contr. info advantage - types of contributions	5,042	2.45	0.75	1	3	2.43	2.36	2.57	2.59	2.49	2.59	2.40	2.41	2.50
Contr. info advantage - share of effort	5,029	2.19	0.77	1	3	2.18	2.15	2.24	2.25	2.30	2.31	2.14	2.21	2.17
Contr. info advantage - importance of contributions	5,030	2.11	0.75	1	3	2.09	2.09	2.20	2.12	2.18	2.20	2.07	2.14	2.07
Contr. info advantage - share of credit	5,014	2.08	0.77	1	3	2.06	2.03	2.18	2.16	2.17	2.20	2.03	2.10	2.06
Contribution statements overall addtl info	4,964	2.39	0.71	1	4	2.38	2.32	2.41	2.50	2.42	2.50	2.34	2.37	2.41
Useful: Share of effort	5,983	2.19	0.83	1	4	2.19	2.19	2.11	2.15	2.24	2.33	2.13	2.23	2.11
Useful: Importance of contributions	5,973	2.24	0.86	1	4	2.23	2.28	2.24	2.06	2.32	2.33	2.20	2.29	2.15
Contributions discussed: All authors	5,305	0.43		0	1	0.41	0.45	0.51	0.58	0.43	0.43	0.44	0.44	0.42
Contributions discussed: Some authors	5,305	0.36		0	1	0.39	0.35	0.29	0.19	0.30	0.34	0.37	0.33	0.40
Contributions discussed: Corresponding only	5,305	0.21		0	1	0.20	0.20	0.20	0.22	0.27	0.23	0.19	0.23	0.18
Contributions approved: All authors	5,376	0.70		0	1	0.69	0.71	0.71	0.77	0.69	0.68	0.71	0.68	0.72
Contributions approved: Some authors	5,376	0.14		0	1	0.15	0.14	0.15	0.10	0.13	0.15	0.14	0.14	0.14
Contributions approved: Corresponding only	5,376	0.16		0	1	0.16	0.15	0.15	0.13	0.18	0.18	0.15	0.17	0.14
Importance of contribution statement	5,701	2.57	0.96	1	4	2.58	2.65	2.43	2.70	2.58	2.85	2.46	2.61	2.52
How common: Senior ghost	5,608	1.97	0.77	1	4	1.98	1.94	1.99	1.95	1.96	1.89	2.01	1.95	2.01
How common: Junior ghost	5,603	2.26	0.79	1	4	2.26	2.29	2.20	2.37	2.28	2.40	2.18	2.28	2.23
How common: Senior guest	5,587	2.67	0.86	1	4	2.67	2.66	2.64	2.79	2.71	2.90	2.56	2.69	2.64
How common: Junior guest	5,591	1.97	0.72	1	4	1.98	1.88	1.96	2.00	1.93	1.90	1.99	1.94	2.00
Postdoc weight: Author position	5,965	3.07	0.78	1	4	3.13	3.04	2.94	2.91	2.94	3.01	3.13	3.02	3.17
Postdoc weight: Contribution statements	5,961	2.59	0.83	1	4	2.57	2.65	2.64	2.63	2.62	2.71	2.54	2.63	2.53
Bio/Life Sciences	6,002	0.65		0	1	1.00	0.00	0.00	0.00	0.00	0.58	0.68	0.60	0.73
Medical/Health Sciences	6,002	0.12		0	1	0.00	1.00	0.00	0.00	0.00	0.13	0.10	0.19	0.00
Physical Sciences	6,002	0.11		0	1	0.00	0.00	1.00	0.00	0.00	0.12	0.11	0.06	0.19
Social Sciences	6,002	0.04		0	1	0.00	0.00	0.00	1.00	0.00	0.06	0.04	0.02	0.08
Other fields	6,002	0.08		0	1	0.00	0.00	0.00	0.00	1.00	0.11	0.06	0.13	0.00
Article age	6,002	369	300	0	1191	394	231	448	428	228	314	399	233	602
Team size	6,002	6.27	3.00	2	14	6.49	6.71	5.69	4.54	5.56	5.55	6.50	6.07	6.60
Position at publication: PhD	5,766	0.13		0	1	0.11	0.19	0.11	0.21	0.20	0.47	0.00	0.17	0.06
Position at publication: Postdoc	5,766	0.14		0	1	0.14	0.12	0.18	0.17	0.17	0.53	0.00	0.14	0.15
Position at publication: Faculty no lab head	5,766	0.17		0	1	0.15	0.27	0.19	0.20	0.19	0.00	0.26	0.20	0.11
Position at publication: Faculty lab head	5,766	0.48		0	1	0.54	0.31	0.46	0.35	0.31	0.00	0.74	0.39	0.62
Position at publication: Other	5,766	0.08		0	1	0.07	0.11	0.07	0.08	0.13	0.00	0.00	0.09	0.06
Remember project well	5,765	3.75	0.47	2	4	3.76	3.71	3.75	3.74	3.72	3.77	3.74	3.73	3.79

table S12. Regression analyses of survey responses on general opinions regarding contribution statements.

	1	2	3	4	5	6	7	8	9	10	11	12	13
	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit	ologit
VARIABLES	adv_type	adv_effort	adv_imp	adv_credit	overallinfo	useful_share	useful_imp	seniorghost	seniorguest	juniorghost	juniorguest	pd_position	pd_statement
Bio/Life Sciences	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
Medical/Health Sciences	-0.113	-0.165	-0.125	-0.181	-0.156	-0.151	-0.071	0.041	-0.118	0.011	-0.195*	0.006	0.043
	(0.099)	(0.094)	(0.098)	(0.095)	(0.099)	(0.078)	(0.080)	(0.087)	(0.086)	(0.089)	(0.084)	(0.084)	(0.078)
Physical Sciences	0.349**	0.090	0.298**	0.271**	-0.017	-0.168*	0.043	0.019	-0.127	-0.168*	-0.033	-0.520**	0.185*
	(0.097)	(0.083)	(0.085)	(0.082)	(0.091)	(0.083)	(0.081)	(0.082)	(0.080)	(0.078)	(0.084)	(0.081)	(0.080)
Social Sciences	0.359*	0.156	0.064	0.241	0.198	-0.137	-0.385**	-0.034	0.124	0.225	0.087	-0.525**	0.124
	(0.152)	(0.137)	(0.122)	(0.135)	(0.129)	(0.141)	(0.124)	(0.117)	(0.123)	(0.126)	(0.137)	(0.128)	(0.131)
Other Fields	0.234	0.211	0.082	0.134	0.056	-0.095	0.021	0.073	-0.072	-0.098	-0.032	-0.217*	-0.029
	(0.121)	(0.112)	(0.105)	(0.104)	(0.117)	(0.090)	(0.092)	(0.096)	(0.096)	(0.102)	(0.093)	(0.099)	(0.092)
PhD	0.547**	0.395**	0.246**	0.408**	0.621**	0.558**	0.397**	-0.356**	0.837**	0.549**	-0.306**	-0.401**	0.534**
	(0.097)	(0.088)	(0.087)	(0.089)	(0.091)	(0.077)	(0.079)	(0.078)	(0.084)	(0.083)	(0.080)	(0.079)	(0.076)
Postdoc	0.553**	0.498**	0.398**	0.441**	0.412**	0.393**	0.235**	-0.325**	0.929**	0.521**	-0.225**	-0.179*	0.306**
	(0.091)	(0.081)	(0.080)	(0.081)	(0.079)	(0.077)	(0.075)	(0.077)	(0.073)	(0.075)	(0.076)	(0.072)	(0.075)
Faculty not lab head	0.061	0.163*	0.245**	0.171*	0.227**	0.163*	0.178*	-0.215**	0.400**	0.173*	-0.195*	-0.185*	0.200**
	(0.082)	(0.079)	(0.079)	(0.076)	(0.085)	(0.069)	(0.070)	(0.076)	(0.076)	(0.075)	(0.076)	(0.075)	(0.072)
Faculty and lab head	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
Other position	0.464**	0.439**	0.332**	0.340**	0.514**	0.004	0.161	-0.120	0.651**	0.592**	0.109	-0.637**	0.280**
	(0.123)	(0.105)	(0.107)	(0.109)	(0.110)	(0.099)	(0.092)	(0.096)	(0.100)	(0.096)	(0.101)	(0.095)	(0.093)
PLOS ONE	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
PNAS	0.285**	-0.043	-0.162**	-0.095	0.171**	-0.208**	-0.241**	0.138*	-0.000	-0.028	0.086	0.350**	-0.170**
	(0.066)	(0.062)	(0.061)	(0.061)	(0.065)	(0.058)	(0.057)	(0.058)	(0.057)	(0.057)	(0.060)	(0.057)	(0.058)
Observations	4,849	4,837	4,838	4,823	4,791	5,757	5,748	5,604	5,583	5,599	5,586	5,735	5,732
chi2	109.9	70.37	61.11	74.08	80.41	99.36	76.27	45.73	233.4	111.0	44.42	167.5	86.29
ll	-4425	-5116	-5153	-5181	-5063	-6896	-7089	-6292	-6815	-6504	-5967	-6413	-6916

Note: Robust standard errors in brackets. *=significant at 5%, **=significant at 1%.

table S13. Regression analyses of survey responses on specific articles.

	1a	1b	2a	2b	3
	mlogit		mlogit		ologit
VARIABLES	discussed_some	discussed_corronly	approved_some	approved_corronly	imp_appearstatement
Bio/Life Sciences	omitted	omitted	omitted	omitted	omitted
Medical/Health Sciences	-0.215*	-0.291*	-0.252	-0.286*	0.003
	(0.108)	(0.125)	(0.135)	(0.130)	(0.082)
Physical Sciences	-0.417**	-0.139	0.138	-0.049	-0.287**
	(0.111)	(0.122)	(0.131)	(0.129)	(0.078)
Social Sciences	-0.726**	-0.100	-0.147	-0.181	0.220
	(0.185)	(0.177)	(0.234)	(0.210)	(0.136)
Other Fields	-0.017	0.180	-0.131	-0.016	-0.116
	(0.131)	(0.132)	(0.160)	(0.140)	(0.094)
PhD	0.190	0.097	0.369**	0.152	1.043**
	(0.106)	(0.121)	(0.126)	(0.124)	(0.082)
Postdoc	0.218*	0.505**	0.156	0.351**	0.630**
	(0.103)	(0.110)	(0.126)	(0.113)	(0.073)
Faculty not lab head	0.076	0.282**	-0.064	0.162	0.285**
	(0.095)	(0.106)	(0.121)	(0.109)	(0.070)
Faculty and lab head	omitted	omitted	omitted	omitted	omitted
Other position	0.045	0.402**	0.364*	0.214	0.256**
	(0.129)	(0.138)	(0.145)	(0.147)	(0.092)
PLOS ONE	omitted	omitted	omitted	omitted	omitted
PNAS	0.401**	-0.250*	-0.052	-0.335**	0.009
	(0.092)	(0.110)	(0.113)	(0.111)	(0.069)
Team size = 2	-3.277**	-0.532**	-2.394**	-0.260	-0.338**
	(0.309)	(0.166)	(0.400)	(0.176)	(0.114)
Team size = 3	-0.949**	-0.139	-0.638**	-0.102	-0.071
	(0.134)	(0.148)	(0.180)	(0.155)	(0.103)
Team size = 4	-0.607**	-0.076	-0.329*	-0.083	-0.090
	(0.126)	(0.147)	(0.162)	(0.152)	(0.092)
Team size = 5	-0.191	-0.024	0.159	0.063	0.001
	(0.123)	(0.151)	(0.152)	(0.151)	(0.097)
Team size = 6	omitted	omitted	omitted	omitted	omitted
Team size = 7	0.162	0.226	0.309	0.146	0.056
	(0.133)	(0.164)	(0.161)	(0.162)	(0.103)
Team size = 8	0.145	0.157	0.299	0.253	0.044
	(0.142)	(0.178)	(0.172)	(0.170)	(0.106)
Team size = 9	0.335*	0.246	0.305	-0.068	0.110
	(0.151)	(0.189)	(0.178)	(0.190)	(0.117)
Team size = 10	0.209	0.273	0.165	0.088	0.147
	(0.161)	(0.198)	(0.197)	(0.198)	(0.128)
Team size = 11	0.236	0.107	0.148	0.032	-0.018
	(0.182)	(0.234)	(0.223)	(0.224)	(0.135)
Team size = 12	0.586**	0.006	0.236	-0.199	0.240
	(0.210)	(0.291)	(0.244)	(0.274)	(0.155)
Team size = 13	0.475*	-0.100	0.466	-0.324	0.247
	(0.229)	(0.326)	(0.258)	(0.328)	(0.169)
Team size = 14	-0.046	0.663*	0.634*	0.608*	-0.083
	(0.286)	(0.304)	(0.308)	(0.300)	(0.248)
Article age	-0.001*	-0.001	-0.001	-0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Article age squared	0.000	0.000*	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Remember project well	-0.040	0.160	-0.162	0.077	0.492**
	(0.069)	(0.082)	(0.083)	(0.085)	(0.047)
Constant	0.278	-1.268**	-0.795*	-1.654**	
	(0.287)	(0.349)	(0.349)	(0.357)	
Observations	5,268	5,268	5,339	5,339	5,694
chi2	427.9	427.9	157.1	157.1	366.5
ll	-5260	-5260	-4276	-4276	-7365

Note: *=significant at 5%, **=significant at 1%. Models 1 - 2 are multinomial logit regressions with “all authors” as reference category (_some, some but not all authors; _corronly, corresponding author only).

table S14. Illustrative responses to the question "Why would you not pay more attention to contribution statements?"

Response	Key issues
It depends on how detailed they are. They are often too generic to be useful. In terms of authorship order, only the first and last author positions really have clear weight. [Faculty lab head; Bio/Life Sciences]	Lack of detail
I think author order is generally interpreted (and should represent) an overall weight of the holistic contribution made to the entirety of the manuscript that contributions don't necessarily capture. Contributions don't necessarily specify the amount of effort between two writers for example - one writer may have contributed 75%, while the other contributed 25%. [PhD; Social Sciences]	Lack of detail
I think the various contributions themselves may be more or less important depending on the study - study design might have been very important on one occasion, and a new application of a statistic method may be important in another where the study design was largely following a well-established protocol. How important these factors were to the success of the research may not be clear to someone not an expert in the field, but sometimes the author order can give you that information - what the authors considered to be the most important part of the contribution. [Postdoc; Earth Sciences]	Lack of detail
If someone is included in a specific aspect of the contribution statement, it is difficult to know *how much* this person was involved in this aspect. Contribution statements are yes/no questions. In addition, and in part because of this yes/no character, different labs write them differently. And sometimes, all authors are included in all categories to avoid conflicts or complications. [Faculty; Bio/Life Sciences]	Lack of detail Low accuracy – social factors
Good question. I think I should. But it's not standard practice across academic journals so the information isn't always available, and when it is, it's not as readily available as simply glancing at author order. [Postdoc; Social Sciences]	Low visibility/costly to access
In part because we don't take the time to look at the paper and/or read contribution statements. When I am reviewing applicants for a faculty search, I usually just read their CV. Too many applicants come in to give the time to finding the paper and reading the statement. Plus, I don't think the contribution states are prepared with care, and thus I don't really trust their accuracy. Only thing I need to know is whether the postdoc is a first author, second author, or middle author. [Faculty lab head; Bio/Life Sciences]	Low visibility/costly to access Low accuracy – lack of care
Because I see author order as more prestigious, and as a clear "ranking" of authors. People tend to take this ranking quite seriously, and I would expect more thought to have gone into it than for the contribution statements, which are not nearly as visible. [Postdoc; Computer Sciences]	Low accuracy – lack of care
I would only use a contribution statement to understand how someone participated in a project that they did not "drive". If the person is first author (or last or corresponding) this typically indicates that they were significantly responsible for the overall idea of the project and central to its success. If the person is a middle author they are likely less central, and so a contribution statement can help clarify what skills or methods that person can bring to someone else's project. More generally, I know that people do spend time	Low informational value Low accuracy – lack of care

thinking about author order because it is a very prominent indicator of contribution. The contribution statement is far less prominent and often filled in by rote during a tedious submission process, and is thus likely to be a less useful signal of contribution. [Faculty lab head; Bio/Life Sciences]	
Most labs that I know are organized in a very hierarchical fashion. Thus, any statement by someone from the middle of the hierarchy ("a postdoctoral researcher") could potentially be either explicitly or implicitly forced by superiors. Generally, I think that determining the performance of scientists from the written text in large collaborative projects with many authors is not a good idea. One way out is to look at the actual work of a person, and to interview the person about her or his work. This can be done, e.g., at conferences. [Postdoc; Physical Sciences]	Low accuracy – social factors Low informational value
1. It takes more effort to go look contribution statements for specific publications, than a simple publication list (e.g. on a CV). 2. contribution statements don't exist for all journals and are only being incorporated more broadly in recent years. 3. Traditionally, I have used author order to indicate the level of participation (regardless of what kind of participation) by each author. [Postdoc; Earth Sciences]	Low visibility/costly to access Inconsistent use by journals Path dependency Lack of detail
It depends probably on the authors but in general to avoid conflicts in a collaboration, specific contributions are often completed not to offend any of the contributors. They do therefore not always reflect reality and contributions are assigned to authors that are not in relation to their efforts (at the expense of the authors that did most). So whereas in the order of the authors this unequal contribution is not being made explicit with the specific acknowledgements it needs to make explicit. So the value of the specific contributions depends on how strongly you want to argue... [Faculty lab head; Bio/Life Sciences]	Low accuracy – social factors
Not all journals require contribution statements and so not all papers, nor all postdocs, could be evaluated in this way. I wouldn't want to use this in evaluations if the information wasn't available for all candidates. [Faculty lab head; Engineering]	Inconsistent use by journals
The standard contribution categories are lacking several very common, but unsavory designations. For example, there are many contributions that help a paper along but are not "intellectual contributions." For example, making a killer figure that increases the profile of the paper, or sharing the statistical software that the author's lab doesn't own. In my own experience, the contribution categories only explain ~66% of the range of efforts that went into the paper. [Faculty lab head; Medical/Health Sciences]	Fail to capture important categories of contributions
Customs for author order are more established (and hence meaningful) than customs for contribution statements. [Faculty lab head; Bio/Life Sciences]	Novelty – lack of established norms
They are relatively new, there are no clear rules on how to write them, and I don't have much experience in using them to judge others. [Faculty lab head; Bio/Life Sciences]	Novelty – lack of experience

table S15. Illustrative responses to the question "Do you have any other comments on this topic that you would like to share? How do you think contribution statements could be improved?"

Specific changes – pros and cons	Key issues
I think contribution statements should be expected for every submission to reinforce standards and expectations for co-authorship and to encourage frank discussion in teams. They are currently not very useful for article readers, and not commonly used to evaluate job candidates - although it would be great if they were common enough to do so. [Staff scientist at National Lab; Physical Sciences]	All journals should use
Suggestions for improvement: make author contributions mandatory for all manuscripts and cross-reference the final contributions to authors ORCID. [Postdoc; Physical Sciences]	All journals should use
I think contribution statements make a good effort at shedding light on a usually opaque and often unfair process. They are a step in the right direction but could be improved by adding a contribution category that would fit many of the honorary senior faculty that end up on these papers because they run the lab, often providing nothing but space or salary. Make a category for "financial supporter" or "lab head". [Faculty lab head; Bio/Life Sciences]	Expand categories
I am very keen to see more detailed author contributions - I'd like see the author order paradigm stopped and replaced with detailed contribution statements. [Faculty lab head; Bio/Life Sciences]	More detail
I think that contribution statements should be much more detailed, in accordance with the specific experiments of each paper. Who contributed to that particular experiment, to what extent? Who analyzed this set of data? etc. Some experiments are more important (or challenging) than others in a paper and this should be acknowledged. Also, contributions statement should be standardized (same format) for all journals so that it can receive more weight for career advancement purpose etc. [Postdoc; Bio/Life Sciences]	More detail Standardize
Contributions statements are frequently an afterthought, and authors don't always agree. This could be mitigated with a culture of making the author contributions more granular, and openness when discussing them. Author contributions can potentially replace the author order if done seriously. I would be happy to see a move towards honest contribution statement and alphabetical author order. [PhD; Bio/Life Sciences]	More detail Culture of open discussion
I think that if co-authors had to 'share the wealth' of a publication, quantified with a listed Author 1 (85%), Author 2 (10%), Author 3 (5%) this would raise the contribution statements to a worthy discussion point. At the moment, it is simply another of the many boxes related to submitting a paper, and from the junior perspective these decisions are made from the position of 'fear of repercussions' rather than truth. [Postdoc; Physical Sciences]	Quantify
I think they are pretty good the way they are. It takes time to establish them, so I wouldn't make it too complicated in the beginning - especially since the process of publishing a paper is very onerous anyway. I like choosing the contributions from the list. I found it very practical and a nice additional information that also gives credit to the right people. Lastly, if people aren't	Keep simple Standardize

honest about it, there is really nothing you can do about it and power play always happens. [Postdoc; Bio/Life Sciences]	
I think contribution statements should be more free form, to describe what someone actually did. The prescribed options are too restrictive, in my opinion. I think that assigning percentages is a horrible idea because parsing effort at that level is completely subjective. [Faculty lab head; Physical Sciences]	Do not standardize Do not quantify
The idea of contribution statements is good. Unfortunately, they do not reflect the actual contributions, not only in PNAS. Senior authors are too dominant in this process. There should be a direct way for all authors to submit their statement to the editor, or have all authors decide what contributions the other authors made. I guess some senior authors will be voted out. [Faculty lab head; Bio/Life Sciences]	Independent submission by individual authors Collective decision making
A problem with contribution statements is that they do not appear in CVs or databases (e.g. PubMed) so are far less visible than author order. For them to be more valuable, a way must be found of making them more prominent. [Faculty lab head; Bio/Life Sciences]	Increase visibility
They are often placed at the end of the paper and might be better placed with the author list. [PhD; Bio/Life Sciences]	Increase visibility
Concerns	
By omitting them. The world is full of grey, not black and white. We discuss ideas as a group, sometimes over a period of years. Ideas emerge from interactions. It is often unclear, and typically unhelpful, to decide who conceived an idea. My students write a first draft, which frequently is a learning experience. I then rewrite the manuscript, often with little or none of the original student content. Who drafted the manuscript? [Faculty lab head; Medical/Health Sciences]	Inability to capture complexity of team work
I am not comfortable at dissecting a team's work. Real Madrid won the Champions league; the contribution of Cristiano Ronaldo was ranked as paramount; however neither Cristiano Ronaldo nor Real Madrid would be praised if the shirts the players wore were not properly washed and were itching them throughout the game. Considering this was Cristiano Ronaldo the reason real Madrid won? [Faculty lab head; Bio/Life Sciences]	Inability to capture complexity of team work
It is a difficult balance. I guess the more we try to specify exactly what everyone does....the more opportunities for conflicts exist. I think rather than trying to find different tools to give credits for authors we should promote "team science". The concept that everyone is actually important to do a good piece of science ("doers", "thinkers", managers, technicians ....and so on). [Faculty lab head; Bio/Life Sciences]	Risk of conflict Inability to capture complexity of team work
I believe contribution statements are a divisive element of the publication process that intensify competition among lab members and produce distortions in behavior as members jockey for position/asserted roles in publication. [Faculty lab head; Bio/Life Sciences]	Risk of conflict
Personally, I think that contribution statements don't make that much difference. All they are doing is formalizing in slightly more detail the same issues as in the order of authors. It seems to me as though people felt that by asking people to write their contributions to a paper, this would resolve disputes about author order. But this misunderstands the problem, in the	Biased assessments of own contributions

main it is not that people are trying to squirrel their way onto papers or higher in the author order than they deserve, it is that people are genuinely oblivious to how much work they have contributed to a paper. If you separately emailed co-authors of a paper and asked the simple question "what percentage of the work did you contribute to this article", the sum of all answers would probably be in the range of 200-500%, depending on the number of authors. Contribution statements won't alter people's inflated perceptions of how much they contributed to a body of work, particularly if they make personal investments into the thoughts and intellectual input into the paper. [Postdoc; Bio/Life Sciences]	
Generally, I'd say the less administrative overhead and the more scientific research we can do as academics, the better. The idea of contribution statements is great, in principle. Yet, I get distracted and annoyed already at the beginning of publishing a paper when I have to submit a contribution statement, a statement of significance, a graphical abstract, an ethics statement, a page charge statement, ... Obviously, not all of those apply to PNAS (and some of those are important), but the key focus should be science, right? [Postdoc; Physical Sciences]	Administrative burden
Perceived benefits	
I think contribution statements are an effective and important tool for clarifying author contributions. Authorship order is often difficult to decide amongst authors, and frequently do not reflect actual contributions, especially when there are large authorship teams -- it is extremely useful to know who carried the work at all stages versus those who simply contributed samples, for example. While I'm sure contribution statements can (and are) manipulated by some, it does at least provide some transparency. I would gladly provide them for all articles that I co-author, except most journals do not request them. I would be glad to see this change! [Postdoc; Social Sciences]	Transparency
I am more and more leaning towards only using journals that ask for these statements because I am sick of people being added "just because". I feel you should work for authorship and with the new generation, they are expecting co-authorship just for showing up. Please do not get rid of this aspect. It is very important. For the next paper I am going to try asking all authors to evaluate everyone else and submit their vote to me as first author and then see how fair the votes are. [Postdoc; Bio/Life Sciences]	Encourage discussion among co-authors
I find the contribution statements to be a valuable teaching tool to help students understand the obligations of authorship. It also gives junior members something to stand on when saying "your contributions thus far do not warrant authorship". [Faculty lab head; Bio/Life Sciences]	Teaching tool Transparency
I really think the "contribution statements" are very important. Being included in a mere list of authors can be justified even with minuscule contributions. Asking your co-authors to accept that you have done something you haven't is much more complicated. Additionally, they provide very useful information for recruiting purposes. [Faculty lab head; Physical Sciences]	Less social influence Informational value

I think they're helpful particularly because there is not the same convention in all fields (e.g. biology and math - last author isn't important in some fields, some list supervisor first regardless etc.) regarding author order so sometimes hard to know what a contribution really involved. [Senior researcher; Bio/Life Sciences]	Consistency across fields
Author order is only an informal way of indication of size and type of contribution and varies per field. Explicit author contributions are formal declarations and make authors consider proper scientific conduct (give credit to whom credit is due) and they can be used in cases of fraud or other retraction issues to hold people accountable. [Faculty lab head; Bio/Life Sciences]	Consistency across fields More accurate credit and accountability for misconduct
I think having contribution statements will greatly reduce ambiguity in authorship. This was my first paper with PLOS ONE, now that I am familiar with contribution standards I will be better able to inform co-authors in advance of the expectations and requirements of authorship. Some conflicts, such as jockeying for last author position by senior team members, will inevitably still occur. [Senior researcher; Bio/Life Sciences]	Tool to plan and agree beforehand
I think current format of statements using the stock categories is not useful at all. The authors should be allowed to state the exact contributions specific to that paper. For example, Author A did experiment A, Author B analyzed data C, Author D developed reagent E, etc. These specifics will allow reviewers, readers, and future employers to infer quickly the expertise of each author. It is not uncommon to attribute most of the contributions to the first author and hiring based on that wrong understanding turned into disaster. It will probably help reduce gift authorships. [Faculty lab head; Bio/Life Sciences]	(If improved:) Information about specific expertise Reduction in gift authorships.

table S16. Authorship positions and contributions controlling for quantity and quality of previous publications.

	1	2	3	4	5	6	7
	Poisson	LPM	LPM	LPM	LPM	LPM	LPM
VARIABLES	i_count contributions	i_conceived	i_performed	i_materials	i_analyzed	i_wrote	i_other
First author	0.242**	0.022**	0.458**	0.006	0.267**	0.094**	-0.034**
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)
Middle author	-0.362**	-0.470**	0.147**	0.043**	-0.165**	-0.474**	0.010**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)
Last author	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_teamsize	-0.032**	-0.017**	0.009*	-0.034**	-0.012**	-0.026**	-0.011**
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
t_teamsize_sq	0.001**	0.000	-0.000	0.002**	0.000	0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
i_lnpriorpubs_quantity	-0.001	0.057**	-0.139**	0.052**	-0.021**	0.038**	0.009**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
i_lnpriorpubs_quality	0.030**	0.017**	0.016**	0.001	0.022**	0.019**	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
t_totalactivitieslisted	0.107**	-0.003	0.013**	0.213**	-0.022**	-0.048**	0.114**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)
t_alphaorder	-0.009	-0.013	0.007	0.011	-0.013	-0.023*	0.010
	(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)	(0.009)
t_affiliations_d	0.008	0.007	-0.054**	0.041**	-0.003	0.034**	-0.003
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.005)
t_published	0.000**	0.000**	0.000	0.000*	0.000**	0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	0.138	0.383**	0.319**	-1.040**	0.349**	0.705**	-0.051
	(0.123)	(0.117)	(0.112)	(0.134)	(0.132)	(0.123)	(0.115)
Observations	73,847	73,847	73,847	73,847	73,847	73,847	73,847
R-squared		0.263	0.259	0.105	0.127	0.318	0.094

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. Models 8-9 use only individuals in teams with <6 members. LPM, linear probability model (OLS).

table S17. Authorship positions and contributions using data from papers in the top 10% of article impact (citations).

	1	2	3	4	5	6	7	8	9	10	Non-alphabetical 11	Alphabetical 12
	Poisson i_count contributions	LPM i_conceived	LPM i_performed	LPM i_materials	LPM i_analyzed	LPM i_wrote	LPM i_other	LPM i_conceived & wrote	LPM i_conceived& wrote& analyzed	LPM i_performed & analyzed	Poisson i_count contributions	Poisson i_count contributions
VARIABLES												
First author	0.219**	-0.060**	0.609**	-0.080**	0.284**	0.032*	-0.035**	-0.012	0.188**	0.606**	0.242**	0.251**
	(0.012)	(0.014)	(0.015)	(0.016)	(0.015)	(0.014)	(0.008)	(0.017)	(0.019)	(0.015)	(0.021)	(0.050)
Middle author	-0.400**	-0.573**	0.275**	-0.013	-0.163**	-0.533**	0.008	-0.609**	-0.444**	0.086**	-0.300**	-0.276**
	(0.012)	(0.012)	(0.014)	(0.014)	(0.014)	(0.013)	(0.007)	(0.013)	(0.015)	(0.013)	(0.023)	(0.091)
Last author	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted	omitted
t_teamsize	-0.050**	-0.027*	-0.006	-0.034*	-0.025	-0.036**	-0.016	-0.034**	-0.026**	-0.024*	-0.039	-0.723**
	(0.012)	(0.011)	(0.012)	(0.013)	(0.014)	(0.013)	(0.011)	(0.010)	(0.009)	(0.012)	(0.080)	(0.233)
t_teamsize_sq	0.002**	0.001	0.000	0.001	0.001	0.001	0.001	0.001*	0.001*	0.001	-0.000	0.110**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)	(0.035)
t_totalactivitieslisted	0.126**	0.011	0.028*	0.220**	-0.017	-0.048**	0.118**	0.008	0.018*	0.033**	0.156**	0.089*
	(0.011)	(0.012)	(0.013)	(0.011)	(0.014)	(0.013)	(0.019)	(0.011)	(0.009)	(0.011)	(0.017)	(0.040)
t_alphaorder	-0.020	-0.021	0.074*	-0.048	0.027	-0.079*	-0.024	-0.033	-0.019	0.065		
	(0.030)	(0.036)	(0.037)	(0.035)	(0.041)	(0.035)	(0.024)	(0.032)	(0.035)	(0.037)		
t_affiliations_d	0.027	0.022	-0.091**	0.059**	0.007	0.046*	0.025	0.022	0.014	-0.029	0.014	0.040
	(0.020)	(0.019)	(0.021)	(0.022)	(0.023)	(0.020)	(0.016)	(0.017)	(0.016)	(0.022)	(0.023)	(0.066)
t_published	0.000**	0.000	0.000	0.000	0.000*	0.000**	-0.000	0.000*	0.000*	0.000	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Field fixed effects	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	-0.377	0.474	0.071	-1.083*	-0.058	0.159	-0.129	0.165	-0.041	-0.359	-1.449*	3.843*
	(0.406)	(0.364)	(0.374)	(0.430)	(0.439)	(0.395)	(0.365)	(0.312)	(0.275)	(0.384)	(0.623)	(1.756)
Observations	8,829	8,829	8,829	8,829	8,829	8,829	8,829	8,829	8,829	8,829	1,438	160
R-squared		0.271	0.137	0.091	0.125	0.312	0.117	0.365	0.346	0.199		

Note: Standard errors clustered by article in brackets. *=significant at 5%, **=significant at 1%. Models 11-12 use only individuals in teams with <6 members. LPM, linear probability model (OLS).