Coauthorship in Physics

Eugen Tarnow

Keywords: authorship, credit, physics, intellectual property

ABSTRACT: In a large and detailed survey on the ethics of scientific coauthorship, members of the American Physical Society (APS) were asked to judge the number of appropriate coauthors on his or her last published paper. Results show that the first or second coauthors are more appropriate than later coauthors about whom there is equal and considerable doubt. The probability of any third and subsequent coauthors being judged as inappropriate is 23% for the APS guideline, 67% for the tighter guideline of the International Committee of Medical Journal Editors, 59% if the guideline requires “direct contributions to scientific discovery or invention”. Only 3% of respondents report having personally rejected an undeserving scientist who expected to be an author on the last published paper. Respondents seem to be divided into two non-overlapping populations—those who report no inappropriate coauthorship and those who have a more graduated view.

INTRODUCTION

Appropriate authorship assignment is of considerable importance to both scientists and to the public. Public recognition of intellectual achievement in terms of authorship is a very strong incentive for scientists to perform their work. Also, knowing who did a particular piece of work is valuable since scientists can contact the appropriate colleague, ask questions and obtain data, and the funding agencies can support more productive scientists and optimize return on investment in the scientific “market”.

Appropriate authorship assignment is also one of the most sensitive issues in science and the study of authorship designation is relatively new and very
controversial. The first survey on actual (rather than hypothetical) authorship practice was performed by Ross Vasta. It was rejected in 1979 by the editor of the *American Psychologist*, William Bevan, then one of the foremost authorities in the scientific community (President of the American Association for the Advancement of Science, President of the American Psychological Association, Editor In Chief of *Science*). He suggested there is a need for intimacy when it comes to matters of authorship.

Vasta studied psychologists who were members of the American Psychological Association and found that 28% of respondents answered yes to the question: “Have you personally ever been involved in a situation where you believe your authorship was not commensurate with your input?” He also found that although ethical guidelines existed at the time, they were not specific and not used by scientists. Honorary authorship (authorship given to colleagues who did not contribute substantially to the research) was considered reasonable by 21% of respondents, and Vasta found this to be uncorrelated with professional age (defined as years since receipt of a doctorate).

Swazey, Anderson and Louis, studied self-reported exposure to a variety of types of misconduct within the preceding five years among university professors and graduate students in four academic fields. They found that inappropriate coauthorship was slightly more frequent than plagiarism and while plagiarism was about three times more likely to be committed by students than by faculty, inappropriate coauthorship was about three times more likely to be committed by faculty than by students. The availability of penalties for plagiarism on university campuses is, however, matched by an almost complete absence of penalties for inappropriate authorship.

Kalichman and Friedman carried out a study further refined by Eastwood, Derish, Leash and Ordway. The latter authors surveyed the overall perception of postdocs regarding inappropriate coauthorship of others and self, and their opinions with regard to appropriate criteria for authorship. They surveyed one thousand postdocs at the University of California, San Francisco, an institution primarily devoted to biomedical research. The authors found that fewer than half of the respondents were familiar with any university, school, laboratory or departmental guidelines for research and publication (such guidelines may or may not exist.) When asked to identify contributions that warrant authorship, nearly half believed that being head of the lab was sufficient for authorship, and slightly fewer believed that obtaining funding warranted authorship. Both views are in opposition to the later, somewhat controversial, “Uniform Requirement for Manuscripts Submitted to Biomedical Journals” (URBJ). The intention to award honorary authorship increased dramatically for those who had firsthand experience with inappropriate coauthorship (either by having been asked to list an undeserving author, having been named as an author together with an undeserving author, or having been unfairly denied authorship). The authors concluded that “despite the respondents’ own standards in this matter, their perception of the actual practice of authorship assignment in the research environment has fostered a willingness to compromise their principles.”

Postdoctoral associates in physics were surveyed in the 1990s. The APS authorship guideline (see Appendix A, question 3), not recognized by 74% of
respondents, was found vague and open to multiple interpretations. Half of respondents thought that obtaining funding qualified for authorship according to the guideline while the other half thought it did not. The probability that an author other than the postdoctoral respondent was judged inappropriate was 17%. In 75% of postdoc-supervisor relationships, authorship criteria had never been discussed: in 61% of relationships the criteria for the postdoc’s authorship were not “clearly agreed upon” and in 70% of the relationships the criteria for designating others as authors were not “clearly agreed upon.”

Yank and Rennie studied authorship in a medical journal that requires identification of the specific contributions of the individual authors, and found that 44% of authors did not qualify for authorship even according to a lenient interpretation of the authorship guidelines of the International Committee of Medical Journal Editors.9

The present article covers all levels of physicists, compares several authorship standards, in particular the APS guideline and URBJ, and obtains other information related to the details of authorship assignment. The current study is the first to compare how perceived inappropriate coauthorship varies with the definition of authorship.

METHOD

The survey (Appendix A, pp. 184-185) was produced using Lotus Domino software. A unique survey document was composed for each scientist and stored on a server. No JavaScript was used to increase the number of browsers that could be used to complete it. The survey consisted of 17 questions, each of which had to be filled out “correctly” (there had to be something in the field and, if a number was requested, it had to be a number). Once all the questions were filled out correctly the survey could be submitted. Upon submission the survey document was randomized and the e-mail identification removed; this prevented anybody (including the surveyor) from identifying the owners of the various surveys. It was not possible to submit duplicate surveys.

A single survey invitation was emailed to 37,000 physicists across the world who were APS members and who had email accounts. About 4000 of those email addresses were erroneous, giving an initial sample size of 33,000. The survey specified that a PhD had to have been obtained or was to be obtained. According to the APS 20% of the members do not have PhDs, lowering the initial sample size to about 27,000. The final return rate was 16% (4302 surveys); 3537 surveys were complete with the exception of “correct” answers to question 8 (765 surveys were incomplete and when “submitted”, the survey left the user with an error message stating that these will not be counted in the results). Responses to question 8 were discarded because of the potential for multiple interpretations. A few surveys (76) that had invalid years for the PhD, gave negative numbers or sought to change the average by using otherwise inappropriate numbers were also discarded. To compare, the return rates for paper-based surveys on authorship topics were: 66% for Vasta’s survey2 (303 surveys), 72% and 59%, for faculty and graduate students respectively for Swazey, Anderson, and
Louis survey\(^4\) (2620 surveys), 33% for the one by Eastwood, Derish, Leash and Ordway\(^6\) (330 surveys), and 34-37% complete surveys for the one by Tarnow\(^8\) (68 complete surveys).

The current return rate is relatively small, though the total number of complete returned surveys is much larger than in previous works. The two main differences between the current survey and the previous surveys are (a) the detailed focus on authorship issues, and (b) that the current survey was done via email. Surveys via the internet give different answers than surveys via other media. Taylor\(^10\) compared internet surveys to telephone surveys. He found that internet surveys capture more information to open-ended questions than do telephone surveys, that it may be more effective in addressing sensitive issues (such as authorship credit), that fewer people pick extremes on multiple choice scales and more pick “don’t knows” and, importantly, that demographic weighting must be used because of differences in online access. The current author is not aware of any studies comparing paper-based surveys with internet surveys.

I did not perform a secondary survey of the non-responders but I examined the quality of the data in two ways. First, I compared the age demographics obtained with publicly available US physics PhD graduation numbers corrected by US 1998 life tables of male death rates (older physicists are overwhelmingly men). The two sets of data are shown in Figure 1. The fitted ratio\(^b\) of members to graduates is 57%. There is excellent qualitative and good quantitative agreement (Figure 2): the age ranges of members in their 50s, 60s and 70s are underweighted by 12%, 6%, and 11%, respectively, the 80s are overweighted by 27%, the 90s are off by less than 1%. These differences may originate in difference in internet access/usage (underweighting the older population), in inaccurate email addresses for the youngest members (who may move around more often) or real factors that may have caused real differences between the PhD graduation rates and APS membership such as job availability upon graduation and immigration. There is no publicly available data on APS membership to use for comparisons.

Second, I studied the time dependence of the survey results: If some particular group of people would not respond to the survey at all, it seems reasonable that there should be some group of respondents with somewhat similar views who would answer the survey but with more hesitation than other respondents and thus make the survey results time dependent. If the survey results are not time dependent, I argue that it is unlikely that it misrepresents such groups. Consider questions 10 and 12: many more respondents than would have been predicted by chance indicated that they wrote down the initial authorship list and that they were the most important contributor on the last published paper. Figures 3 (a) and (b) show the moving average over the previous 100 incoming surveys of the answers to the two questions. The result seems to suggest that respondent time is not strongly tied to the answer to the two questions and is thus consistent with the view that the sample is representative of the APS membership.

\(^b\) The fitted ratio was determined by using the age trend of the membership and the exact number of people graduating. The two trends overlap when the estimated probability that a person graduating with a PhD in physics would then become a member of the APS is 57%. That guess is the fitted ratio.
Some results are insensitive to over- and underweighting of the membership demographics. In the survey three measures, seniority, prestige and power, reflect professional age (year in which the respondent earned a PhD), whether the respondent is a “Fellow” of the APS, and how comfortable the respondent indicated he or she would be in denying authorship to an undeserving researcher. These three measures are well correlated (see Figures 4 and 5; more about this correlation below). When studying seniority, the responses of those who obtained their PhD in the year 2001 or later were not included. (I apologize to the ignored respondents).

The statistics reported here refer to papers which “report a discovery or invention” unless specified otherwise.

RESULTS AND DISCUSSION

Appropriateness of Actual Authorship
Three conscious estimates of appropriateness of stated authorship were used. Two authoritative authorship definitions included the APS guideline (question 3) and URBJ (question 4). A third definition limited authorship to those authors who contributed “directly” to the “scientific discovery or invention” (question 7. The phrase in quotation marks is taken from the Nobel definition of why Nobelists are awarded prizes). This question did not appear in a definition format.

Figure 6 shows the number of inappropriate coauthors (papers with 10 or more coauthors were combined into bins with at least 30 respondents). The curves were fit starting with the function

\[ \text{probability that the nth coauthor is inappropriate} = a \cdot (1 - e^{-n/b}) \]  

(1)

where a and b are the fitting parameters. For n>b the probability that the nth coauthor is inappropriate becomes a and for b>n>=0 the probability is somewhere between 0 and a. Using (1), the functional form for the percentage of inappropriate coauthorship shown in Figure 6 is

\[ \text{percentage of inappropriate coauthorship} = \frac{a \cdot (1 - e^{-1/b}) \cdot (1 - e^{-n/b})}{1 - e^{-1/b}} \frac{1}{n} \]  

(2)

The values of a (b) are 23% (2.6) for the APS guideline, 67% (3.8) for URBJ, and 59% (2.1) for direct contributions.

This is the most remarkable result in this paper. First it shows that, qualitatively, inappropriate coauthorship can be well described over almost three orders of magnitude of coauthors with a simple functional form independent of the authorship criteria. Second, the two parameter fit shows that within the same criterion there is apparently little difference in authorship assignment on papers with 4 coauthors or papers with 600 coauthors; the first and second coauthors are typically more appropriate. Thirdly, the figure shows that inappropriate coauthorship is strongly dependent on the criterion used.

The Dynamics
The APS authorship guideline is the only written criterion for authorship specifically for papers in physics. Respondents reported that it was not used in 92% of publishing
experiences (question 15). Excluding the “don’t knows”, 83% of all respondents, report that THE PAPER had a final list of authors that did not change from the initial list (question 11). This was true whether or not the respondent developed that initial list (question 12). In 12% of the cases the final authorship list was longer than the initial list, in 4% of the cases the list was shorter (question 11). Thirty-four percent of respondents reported that they were “not comfortable” denying an undeserving author (question 13), 42% were “somewhat comfortable” and 25% were “comfortable”. These responses varied with age as shown in Figure 5. The largest difference in responses were the “comfortable” responses: 45% for older respondents and only 21% for younger ones. Correspondingly, only 20% of older respondents were “not comfortable” while 39% of younger ones were not; 3% reported denying authorship to a scientist who expected authorship but who did not qualify (question 14).

The new finding that 92% of published papers do not use the APS guideline to determine authorship is not inconsistent with an earlier finding that the majority of postdocs (74%) do not recall seeing the APS statement. That the authorship list stays relatively unchanged (no change in 4 out of 5 publishing experiences) is also not inconsistent with the previous finding of authorship discussions being rare since change could reasonably be expected to be precipitated by discussions. There is an apparent discrepancy between attitude and behavior in terms of denying authorship to an undeserving author, 25% report themselves being “comfortable” about it but only 3% actually did so. This may be because only 3% had occasion to do so since the social penalty for removing an undeserving author can be severe. The strong correlation of professional age with comfort in denying undeserved authorship suggests that the perceived ability to deny authorship to undeserving scientists is strongly correlated with the respect and power that come with age.

Is It Possible to Tell Who Contributed The Most?
When attempting to judge the meritocracy of authorship, one measure is whether a peer can tell from the list of authors who made the greatest contribution to the paper when contributions are unequal (question 9). The term “peer” was used to control for unwritten authorship conventions that particular sub-fields may use (head of the lab goes first, head of the lab goes last, most important contributor goes first, etc.)—a peer is expected to know such conventions. For papers with at least one coauthor, the respondents answered “No” 46% of the time. Thus it is not generally possible for a peer to determine who contributed the most from the information currently present in the byline.

The Number of Authors per Authorship Experience
Figure 7 shows the number of respondents as a function of the number of authors on THE PAPER (question 2). The range of authorship is almost three orders of magnitude; the most common number of coauthors is 1 or 2; the median number of coauthors per authorship experience is 3; the average, dominated by high-author papers, is 15.
Coauthorship in Physics

Definition Preference, Solicited Opinion about the Definitions and a Composite Preferred Definition
The survey asked which authorship definition the respondents preferred. The APS guideline was preferred by 64% (no large correlation with professional age), URBJ by 25%, 6% preferred different criteria and 6% preferred not to have any requirements.

The specific reasons mentioned for preference for the APS guideline were, in order of frequency: URBJ is too restrictive (32% of responses distinguishing the two definitions), the URBJ writing requirement is too restrictive (13%), the URBJ writing and approval requirement is too restrictive (12%), the need for authorship incentives for collaborations (11%), URBJ final approval is too restrictive (9%), etc. The specific reasons mentioned for URBJ preference were, in order of frequency: the APS guideline is not precise or restrictive enough (34%), URBJ lessens abusive authorship practices (32%), involvement in reviewing is needed (10%), involvement in writing and reviewing is needed (8%), URBJ is more ethical (7%), etc. That the APS guideline is not thought to be sufficiently restrictive is consistent with the previous finding that it allows broad interpretation—for example, as mentioned above, it does not clearly indicate whether obtaining funding for a research project qualifies a person for attribution as author.

The opinions of the respondents on the APS and URBJ definitions can be merged to construct a definition that likely would have a higher approval rating than the APS guideline. It would adapt the APS guideline by adding the URBJ restrictive caveats “acquisition of funding, the collection of data, or general supervision of the research group, by themselves, do not justify authorship,” (to please those who preferred the URBJ definition because they thought it decreased abusive authorship practices and was more ethical), would not require actual participation in writing (found to be too restrictive by those who preferred the APS guideline) and would not require approval of the final manuscript (found to be too restrictive by those who preferred the APS guideline), BUT would require a statement on record about the paper from each author. This merged authorship guideline would read:

“Authorship should be limited to those who have made a significant contribution to the concept, design, execution or interpretation of the research study. All those who have made significant contributions should be offered the opportunity to be listed as authors. Other individuals who have contributed to the study should be acknowledged, but not identified as authors. Acquisition of funding, the collection of data, or general supervision of the research group, by themselves, do not justify authorship. A statement from each author on the final version of the paper indicating that that version has been read and whether the author approves or rejects the contents of the paper must be recorded.”

Two Types of Respondents
There is anecdotal evidence for a significant division in the scientific community in terms of authorship issues: some believe authorship is an important issue while others do not. Ross Vasta had some respondents ask him why he cared, and others thanked him for his study. My own experience in authorship research is similar: some
respondents care deeply about fairness in authorship assignment, others could not be bothered and even showed hostility to discussing the topic of authorship assignment.

Figure 8 shows the number of reports of multi-author papers (more than 20 authors per paper) and the number of papers that did not meet the criteria presented in survey questions 3 (APS guideline), 4 (URBJ), and 7 (being a direct contributor to the discovery or invention) in the questionnaire. The majority opinion shifts as the authorship definition changes and there is evidence of two separate populations emerging for respondents to questions 4 and 7 with maxima at the two extremes.

The same data for papers with 5-10 authors is shown in Figure 9. As before, the majority opinion shifts as the authorship definition changes and there is evidence of two separate populations emerging for respondents to questions 4 and 7 (examine the results from each authorship guideline separately). In contrast to Figure 8, respondents who indicate a large amount of inappropriate coauthorship are not at the extreme end, but rather somewhere in the middle.

The distributions presented in Figures 8 and 9 describe opposing perceptions: that stated authorship rather completely fulfills various criteria, and another that, to a large extent, authorship does not fulfill these various criteria.

SUMMARY

This study suggests:

1. Using a two parameter model, inappropriate coauthorship in physics can be well quantified over almost three orders of magnitude in the number of coauthors. There is little quantitative difference in how authorship is assigned whether a paper contains four or 600 coauthors. Papers with three or fewer coauthors are less likely to have inappropriate authors.
2. The probability that an additional coauthor is inappropriate varies strongly with the definition of authorship from 23% (APS guideline) to 67% (URBJ).
3. Forty-six percent of respondents report that the most important contributor cannot be identified from the authorship list.
4. When asked which authorship definition is preferred, the APS guideline is favored by 65% of respondents while the tighter medical editor’s guideline is preferred by 25%. A combined definition that might have a higher approval rating is proposed.
5. Authorship assignment is not very flexible—once the initial authorship list has been determined it typically does not change. There is only a 4% probability of a decrease and a 12% probability of an increase in the number of authors.
6. Survey responses identified two “types” among the respondents: those who believe there is no inappropriate coauthorship and those who have a more graduated point of view.

Since authorship is such an important part of the scientific endeavor, one must ask the question—are there useful ways to standardize the procedure of authorship assignment? There seems to be at least two options. One is to follow the patent authorship model and have an attorney, or another disinterested party, inquire into the
Coauthorship in Physics

research work and, according to existing legal standards for patent authorship, write down the list of authors. A second choice would be to more accurately assign authorship by adding an authorship section at the end of each paper explaining what each author contributed.\textsuperscript{11-12} A non-committal endorsement of this latter option is described in the URBJ guidelines: “Editors may ask authors to describe what each contributed; this information may be published.” A sample “manuscript authorship form” has been previously published.\textsuperscript{8}

Acknowledgment: I would like to extend my gratitude for the time given to me by the survey respondents; for help with construction of survey questions and critical readings of the manuscript by Charles DeLeone, Michelle Fine and Rafi Kleiman, to Mark Hybertsen for pointing out a software error, to Charles Kittel for pointing out the problem with Question 8 and to the APS for assisting me with the survey.

REFERENCES

APPENDIX A: QUESTIONNAIRE

Your thoughts on the authorship requirements of the APS

1. In which year did you obtain, or do you expect to obtain, your doctorate (please use four digits for the year)?

In this questionnaire, THE PAPER will refer to the most recent paper published in a refereed journal with yourself as an author.

2. How many authors are there in total on THE PAPER?

3. In your opinion, how many of the authors on THE PAPER fulfill the requirements for authorship as defined by the APS:
   “Authorship should be limited to those who have made a significant contribution to the concept, design, execution or interpretation of the research study. All those who have made significant contributions should be offered the opportunity to be listed as authors. Other individuals who have contributed to the study should be acknowledged, but not identified as authors.”

4. In your opinion, how many of the authors on THE PAPER fulfill the requirements for authorship as defined by the BIOMEDICAL JOURNAL EDITORS:
   “Authorship credit should be based only on 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published. Conditions 1, 2, and 3 must all be met. Acquisition of funding, the collection of data, or general supervision of the research group, by themselves, do not justify authorship.”

5. Which authorship requirements do you prefer?
   American Physical Society
   Biomedical Journal Editors
   I would prefer different requirements
   I prefer not to have requirements at all

6. Please explain your answer to the previous question.

7. In your opinion, how many of the authors on THE PAPER contributed DIRECTLY to the scientific discovery or invention presented? (If THE PAPER did not report a discovery or invention, please enter 0).

8 (a)-(d). In your opinion, how many authors on THE PAPER contributed ...
   (a) a very significant amount of relevant scientific work
   (b) a significant amount of relevant scientific work
   (c) a somewhat significant amount of relevant scientific work
   (d) an insignificant amount of relevant scientific work
9. Would a peer be able to tell from the authorship listing on THE PAPER which author(s), in your opinion, made the most significant contribution(s)?
   Yes  No

10. Were you the most significant contributor on THE PAPER?
    Yes  No

11. How did the final authorship list change from the initial version?
    The final authorship list was longer than the initial list
    The final authorship list was the same as the initial list
    The final authorship list was shorter than the initial list
    I don’t know.

12. Were you the person who wrote down the initial version of the authorship list on THE PAPER?
    Yes  No

13. Would you feel comfortable denying authorship to a scientist who expects authorship but who, in your opinion, does not qualify for authorship?
    Comfortable
    Somewhat comfortable
    Not comfortable

14. On THE PAPER, did you deny authorship to a scientist who expected authorship but who, in your opinion, did not qualify for authorship?
    Yes  No

15. On THE PAPER, were the APS authorship requirements discussed among potential authors and then used as the agreed-upon criteria used for authorship?
    Yes  No

16. Are you aware that the APS authorship requirements changed in 2000?
    Yes  No

17. Are you an APS Fellow?
    Yes  No

This survey is conducted by Eugen Tarnow, PhD. He is a member of the APS and the Council of Science Editors. He has previously published articles on authorship among physicists in Science and Engineering Ethics and Nature. The results of this survey will be submitted for publication in a refereed journal. Any questions about this survey should be sent to etarnow@avabiz.com

Copyright © 2001 Eugen Tarnow. All rights reserved. Reprinted by permission.
Figure 1. Quality assessment of respondent demographics. The filled squares show the sample population and the unfilled squares show the life-table adjusted physics graduation rates in the United States.

Figure 2. Quality assessment of respondent demographics by decade. The squares show the sample population and triangles show the integrated life-table adjusted physics graduation rates in the United States.
Figures 3 (a) and (b). Time dependence of the running averages of answers to survey questions 10 and 12.
Figure 4. The probability of respondents claiming to be a “Fellow” of the APS vs. professional age.

![Probability of Fellow distinction vs. Decade awarded PhD](image)

Figure 5. Graph showing percentage of respondents who assert he or she would be comfortable, somewhat comfortable or uncomfortable, denying “authorship to a scientist who expects authorship but who, in [his or her] opinion, does not qualify for authorship” versus professional age.

![Percentage of respondents](image)
Figure 6. Inappropriate authorship on all papers reporting discoveries measured three different ways. APS guideline is shown in filled squares and dashed line, URBJ in filled circles and full line, the direct criterion in filled upward-pointing triangles and dash-dotted line.

Figure 7. Number of respondents with $N>0$ coauthors on “The Paper”.

Science and Engineering Ethics, Volume 8, Issue 2, 2002
Figure 8. Number of reports of multiauthor papers (more than 20 authors per paper, a total of 163 responses) and the percentage of coauthors who did not clearly fulfill criteria described in questions 3 (APS – “SOCIETY”), 4 (URBJ – “BIO”), and 7 (being a direct contributor to the discovery or invention). The majority opinion shifts as the authorship definition changes and there is evidence of two separate populations emerging for respondents to questions 4 and 7.

Figure 9. Number of reports of multiauthor papers (5-10 coauthors per paper, a total of 374 responses) and the percentage of coauthors who did not clearly fulfill criteria described in questions 3 (APS – “SOCIETY”), 4 (URBJ – “BIO”), and 7 (being a direct contributor to the discovery or invention). The majority opinion shifts as the authorship definition changes and there is evidence of two separate populations emerging for respondents to questions 4 and 7. In contrast to Figure 8, respondents who indicate a large amount of inappropriate authorship are not at the extreme end, but rather somewhere in the middle.