

THE INTERNATIONAL STUDIES PROFESSION

Symposium on Replication in International Studies Research

The following symposium tackles an important debate in the field of international studies research and in social science research more broadly. Originating as presentations at the 2002 International Studies Association meetings in New Orleans, the following set of papers examines issues concerning the potential replication of research results from a number of different conceptual and technical perspectives. It also spans an array of journals and journal editors in our field that confront issues of replication on a regular basis. The interactions generated by this symposium have already led to the agreement, detailed at the end of the symposium among the four leading international relations journals to adopt a single common replication policy. The editors of these journals challenge others to follow their lead. The Editors of ISP hope that this symposium begins a larger discussion on these issues and invites commentary from interested scholars. Lastly, I want to thank Nils Petter Gleditsch for soliciting, organizing, and coordinating the contributions to this symposium. Without his work, this project would not have come to fruition as quickly or in as insightful a fashion. Three anonymous reviewers also provided detailed comments on the entire set of articles in this symposium. We are indebted to the care and effort they gave to the project.

Mark A. Boyer

For the Editors of ISP

The Replication Debate¹

In September 1995 *PS: Political Science and Politics* published a symposium on replication in the social sciences. Gary King issued a strong call for a replication standard. For any empirical work enough information should be made available so that a third party can comprehend, analyze, replicate, and evaluate the results without further information from the author. Such a policy would allow for further development of the social sciences, he argued. “The most common and scientifically productive method of building on existing research is to replicate an existing finding—to follow the precise path taken by a previous researcher, and then improve on the data or methodology in one way or another” (King, 1995:445). A replication standard would become a pedagogical tool.

Several participants in the symposium shared King’s perspective. Kenneth J. Meir and Linda Fowler, for example, claimed that a data-replication policy would enhance the research process, specifically the methodological standards. Furthermore, important benefits to the field of political science would follow, by facilitating the creation of a community of scholars who would be able to develop a comprehensive body of research based on similar methods and data. “...we all share a common

¹Claire Metelits collected most of the information used in this study. Her work was supported financially by the University of Denver’s Social Science Foundation and the Graduate School of International Studies. The articles in this symposium were first presented to a session on “Replication in International Relations” at the 43rd Annual Convention of the International Studies Association, New Orleans, LA, 24–27 March 2002. We are grateful to the participants at this session, as well as the *ISP*’s editors and reviewers for helpful comments. We would also like to thank Lars Wilhelmsen and Naima Mouhlebl, both at PRIO, for technical assistance.

interest in furthering knowledge,” Meir wrote, and “gains in knowledge occur only with the analysis and scholarly debate” (Meir, 1995:456; see also Fowler, 1995).

Other participants in the symposium argued strongly *against* enforcing a data replication standard. For example, James Gibson wondered what such a standard would do to the next generation of scholars. A standard that was applied across the entire field of political science would see a vast expansion of strong data sets. Yet the result, he argued, would be counterproductive. “Generations of scholars will be raised to mine existing data sets, not to collect data on their own. Data hawks will be replaced by data vultures.” He also feared that a replication standard might engender a “belief systemization” of the field in which scholars witness the production of countless papers that discuss minute methodological issues. Furthermore, the users of easily accessible data sets would not be held to the same standards as those who produce original data (Gibson, 1995:475).

Indeed, numerous issues arise when the social science community has to deal with data sharing, as Sieber (1991a) attests in a book dedicated to this subject. These issues include the condition and documentation of the data set, as well as the effort and costs of data sharing. A further issue is the qualifications of those researchers who are requesting the data (Sieber, 1991a:3).

Sieber describes some of these concerns that circulate in the social science community as “naïve” or “informal.” The naïve concerns, usually made by neophyte sharers, include the following: who owns the data; the cost of documentation; whether the researcher will receive professional credit (i.e., tenure); whether the researcher will be cited in the literature for sharing the data; whether the data that stems from funded research is controlled by the investigator, the institution with which the investigator is affiliated, or the funding agency; in collaborative projects, who the responsible party will be for orchestrating the sharing of data and answering questions and criticisms (Sieber, 1991b:142–145).

Yet these concerns lead to the more significant issues surrounding the data-replication debate. What data, for example, should not be shared? Documentation, for instance, is often thought of as a waste of time. However, secondary users of data do not often possess the intimate knowledge sufficient for replication. What kind of data should researchers archive? Finally, data sharing offers unique and enriching opportunities for cross-disciplinary contributions (i.e., political science questions may be answered using data gathered by economists). These opportunities also bring with them the dilemma of using data that have been collected and “cleaned” in ways that are unfamiliar to the secondary user (Sieber, 1991b:145–149).

Replication Policies Today

In the seven years that have passed since the publication of the symposium in *PS*, a number of journals have introduced replication policies. We have looked systematically at the 15 most frequently cited journals in political science and the 13 most frequently cited journals in international studies,² according to the 1999 edition of *Journal Citation Reports*.³ *Journal of Conflict Resolution* appears on both lists, so we have looked at 27 journals altogether. The data on their replication policies were collected between 15 June and 27 July 2001.

We first inspected each journal’s website. If information on its replication policy was available on the Web, it was usually found on the page directed at prospective authors. If the information was not available there, we attempted to obtain a copy of

²Two international law journals, *American Journal of International Law* and *Ocean Development and International Law*, were among the top 15 international relations journals, but were omitted from our survey.

³*Journal Citation Reports, Social Sciences Edition* is published by the Institute for Scientific Information. It summarizes the citations published in *Social Sciences Citation Index* and ranks journals by their “impact factor.” See www.isinet.com/isi/products/citation/jcr/.

a recent issue of the journal, since similar information is frequently found in the back of the journal or on the inside cover. Finally, if the policy was still unclear, we personally contacted the editorial office, explaining the purpose of our study. In the case of non-response, we sent two reminders.

Few journals had a replication policy clearly stated on the webpage or in the instructions printed in the journal itself. Therefore, we obtained most of this information by e-mail from the editors. On a few occasions, there was confusion as to what “data replication” meant. Some editors thought we were referring to copyright policies. When this matter had been clarified, these editors frequently did not respond at all. We took this to mean that the journal does not publish articles that contain quantitative data or large-*n* data sets, and thus have no urgent need for considering a replication policy.

Table 1 summarizes the results of our survey of whether the journals have a stated replication policy. Slightly less than one-fifth of the journals do. If we include the journals that do not have a stated policy, but say that they encourage replication, we end up slightly short of one-half. Overall, the international relations journals appear to be slightly ahead of the political science journals. Among the leading political science journals *American Political Science Review* “requires” replication, while *American Journal of Political Science* merely “encourages” it. Among the international relations journals, *International Studies Quarterly* “requires,” *International Organization* “encourages,” while *World Politics* “expects” readers to have access to data used by its authors. For a more detailed report see the Appendix.

Even for journals that do have a stated policy, the practice may vary. In the following section we review in greater detail the policies and practices of four journals in international relations where a large proportion of the articles have quantitative data, as well as two political science journals with a similar orientation and which are seen as important outlets for work in international relations.

Practices

We discuss the six journals in alphabetical order. All of them, except *AJPS*, claim to have a replication policy, and *AJPS* is among those that encourage replication. In addition to investigating the policy, we checked all the articles in the past five volumes (1997–2001) to see if they generally complied with the journal’s stated policy.

American Journal of Political Science

AJPS encourages authors of articles with quantitative data to include a footnote describing how others can obtain their data and provide documentation of the statistical analysis in order to make it possible to replicate the research. Authors are

TABLE 1. Replication Policies of the Leading Political Science and International Relations Journals, Summer 2001 (%)

Policy	Political Science	International Relations	All Journals
Stated replication policy	13	23	18
No stated policy but encourages replication	27	31	29
Explicitly states it has no replication policy	40	31	36
No response	20	15	18
Total	100	100	100
(n)	15	13	27

Note: *Journal of Conflict Resolution* is included in both the first two columns but only once in the third. For details of the survey see the Appendix.

encouraged to share their data by depositing them in a public archive. The data necessary for replication are often said to be available from the authors. However, e-mail addresses are not always provided and in several cases there was no information at all regarding data location.

American Political Science Review

APSR requires authors of quantitative manuscripts to describe their procedures in enough detail to allow readers to evaluate what has been done. *APSR* provides investigators with the authors' e-mail addresses, though it does not specifically state: "Data available at ..." or "Data available by contacting ..." Less than five articles list the specific locations of the data for replication. There are also recent cases (see 94(2):407) where the authors use quantitative research, but do not make any reference to data replication, though they describe their procedures. This appears to follow *APSR* policy, which only requires authors of quantitative articles to describe their procedures in enough detail that will allow readers to *evaluate* the methods carried out. In some cases, too, the journal provides the name of a specific institute where data are stored, but no direction on how to obtain the information necessary for replication.

International Interactions

II asks authors of data-based articles to indicate how scholars can obtain their data, in addition to other information necessary to replicate the reported results, such as in public archives or with the individual authors (<http://www.gbhap.com/journals/197/197-nfc.htm> as accessed 7 August 2001). This journal and its authors do not appear to follow through on the stated data replication policy. Generally there are few data replication references. Recently, the editor has taken steps to improve the practice (Ray and Valeriano, 2003).

International Studies Quarterly

ISQ requires authors whose articles are accepted for publication to ensure that their data are accessible. This journal provides authors' e-mail addresses at times. What we find most often, though, is a reference to obtaining information from the author without specific contact information. Otherwise there is no reference to the location of data. There are a few (less than 10) instances when the authors provide a Web address that contains the information necessary for data replication. For an update on the *ISQ* practice see James (2003).

Journal of Conflict Resolution

JCR encourages authors of articles using databases to alert readers to the location of the original data, in addition to other pertinent information such as original data, codebooks, etc., which are necessary to replicate the published results. Most of the authors in this journal who use quantitative data do not provide specific information to readers concerning the location of their original data. Very few authors offer a specific Web address where readers can find the data stored. Few provide contact information. In a few instances, references to data are buried in footnotes in the middle of the article. In some cases, the authors list the archive where the information is filed, but there is no specific Web address where a researcher can locate this information. There are also a few instances in which authors claim that data for replication purposes can be obtained from them, yet they do not provide contact information. Subsequent to our survey, *JCR*

strengthened its replication requirement as of the February 2002 issue (Russett, 2003).

Journal of Peace Research

JPR has had a data replication policy since 1998. Authors of published articles with systematic data must post the data, codebook, and other appropriate material on their own websites. If they do not have a website, the information necessary for replication purposes is posted on the *JPR* data replication webpage. A footnote linked to an asterisk at the end of the title must indicate where the data can be found. In general, *JPR* has followed this policy. The journal updates the datapage on a continuing basis. There are quite a few exceptions where Web addresses provided by authors are no longer functioning, or when an author is reluctant to provide the original data. These are followed up from time to time.

Summing Up

Overall, the survey does not seem to provide much encouragement for those who favor a strong replication requirement. Data-replication policies seem to have been strengthened only slightly over the past five years. Journals that proclaim a replication policy often fail to implement or enforce it.

With the widespread use of the Internet, more authors use their e-mail addresses as contact information. But this does not in itself ensure that the data become fully accessible. Several of these journals may have assumed that an e-mail address or the author's institutional affiliation would be sufficient to locate the data. In attempting to access data in this way a researcher wishing to replicate is nevertheless at the mercy of the principal investigator.

In many instances, authors of these journals provided the name of the archive where the data are housed. This may be sufficient for those who are familiar with accessing information from these databanks. However, to the novice researcher, it may appear to be a daunting task, costs money, and is not always user-friendly.

However, the situation may be changing. To some extent, this survey itself has already served as a reminder to the research community of the importance of these issues. Several journal editors have responded that they are considering formulating a replication policy and the editors represented here have strengthened their journals' policies and practices. The Virtual Data Center promoted by King (2003) promises a great step forward in making replication feasible. And the suggestion made by Bueno de Mesquita (2003), to make replication data available to an article's reviewers, will certainly provoke debate.

Nils Petter Gleditsch

*International Peace Research Institute,
Oslo (PRIO) and Norwegian University of
Science and Technology (NTNU)*

Claire Metelits

Northwestern University

Appendix: Journal Policies

We collected data on the replication policies and practices of 27 leading international relations and political science journals between 15 June and 27 July 2001. The journals are discussed in alphabetical order.

International Relations Journals

European Journal of International Relations is a journal for which, according to editor Friedrich Kratochwil, the issue of replication seldom arises. Most of its articles are conceptual in nature. Nevertheless, as a policy they encourage authors to provide the data set to others on request. Often, the mere indication of the data source (such as an available data set or a data bank) provides enough information to allow for a personal check on the data and/or replication of the study (e-mail 11 July 2001).

Foreign Affairs rarely publishes articles that make use of data sets. However, the managing editor claims that if they did, the editors of the journal would insist that authors make them available to the research community for replication purposes (e-mail 13 July 2001).

Foreign Policy does not offer its readers original data in the form of “space-consuming charts and tables”; it provides links to the data on its website. The journal is not refereed and caters to a diverse audience, most of whom, according to James Gibney, managing editor, do not have a need to access original research (e-mail 5 July 2001).

International Affairs (London) The editor was somewhat confused as to the nature of our query. We take this to imply that *International Affairs* does not have a data replication policy, which seems consistent with the nature of the journal.

International Organization does not require authors of published articles to provide access to their data sets, although it encourages them to do so, especially when the data are not yet available to the public. If an author’s data are archived with the Publication-related Archive of ICPSR, or on the author’s own website, *IO* encourages a brief reference to this in a note accompanied by a complete reference listing (<http://mitpress.mit.edu/journal-submissions-guidelines.tcl?issn=00208183> as accessed 3 July 2001).

International Security does not have a formal data replication policy. Many of the journal’s articles do not utilize databases, and when they do, the information is often based upon easily accessible research. Philosophically, according to Steven Miller, editor, it “favors openness and, if asked, would encourage authors to facilitate replication” (e-mail 5 July 2001).

International Studies Quarterly requires authors whose articles are accepted for publication to ensure that their data are accessible. If their data are not already publicly archived, the authors are required to certify that their data will be made readily available to others (<http://www.public.iastate.edu/~isq/> as accessed 3 July 2001).

Journal of Common Market Studies does not have a data replication policy (e-mail 9 July 2001 from Iain Begg, co-editor).

Journal of Conflict Resolution asks the authors of articles using databases to alert readers to the location of the original data, in addition to other pertinent information necessary to replicate the published results. This information should include such items as original data, specialized computer programs, extracts of existing data files, codebooks, and a file that explains how to reproduce the results found in the article. Authors are asked to remove confidential information from their data sets, such as the identity of survey respondents. The journal notes that authors have the right to first publication and may wish to embargo data before they finish publishing from them. However, these authors are required to specify a date by which the replication data set will be made available (<http://kokopeli.library.yale.edu/un/un2fla4.htm> as accessed 3 July 2001). Subsequent to our survey, *JCR* strengthened its replication requirement as of the February 2002 issue (Russett, 2003).

Journal of Peace Research has had a data replication policy since 1998. Authors of published articles with systematic data must post the data, codebook, and the “do,”

“log,” “batch,” or “run” file on their own if the Web does not have one. A footnote linked to an asterisk at the end of the title must indicate where the data can be found (*Notes to the Authors*, version 3 July 2001).

Review of International Studies did not reply to our query. Given the nature of the journal, it seems most likely that it does not have a replication policy.

Security Studies does not have a formal data replication policy. According to Benjamin Frankel, editor, the journal does not publish quantitative or large-*n* studies, so that in the ten years of the journal's existence, this particular issue has not arisen (e-mail 5 July 2001).

World Politics does not have a specific data replication policy. However, the editor indicated that it is expected that readers should have access to data used by the journal's authors. References to data are indicated either within an article itself as an appendix, for example, or by identifying the specific Web address where readers can access the material, or through a statement claiming that necessary materials are available by contacting the author(s) (e-mail 20 July 2001).

Political Science Journals

American Journal of Political Science (AJPS) encourages authors of articles with quantitative data to include a footnote describing how others can obtain their data and providing documentation of the statistical analysis in order to make it possible to replicate the research. Authors are encouraged to share their data by depositing them in a public archive. The data necessary for replication are supposed to be available from the authors (see <http://ajps.org/frames-con.htm> accessed 3 July 2001). However, e-mail addresses are not always provided and in several cases there was no information at all regarding data location.

American Political Science Review requires authors of manuscripts containing quantitative evidence and analysis to describe their empirical procedures in sufficient detail to permit reviewers to evaluate what has been done and other scholars to carry out similar analyses on other data sets. Descriptions of experiments must be in full, including protocols, methods of subject recruitment and selection, payment and debriefing procedures, as well as other relevant information. Authors should not refer readers to other works for similar descriptions of their research methods (<http://www.ssc.msu.edu/~apsr/instruct.html> as accessed 3 July 2001).

British Journal of Political Science requires authors who include statistical analysis of data in their articles to deposit a replication data set in a “major national data archive” (e-mail 11 July 2001).

Comparative Political Studies does not have a formal replication policy. The editor states that the issue of data replication rarely arises. He also states that the journal staff do ask that authors make their data available to readers upon request, and may at times ask authors to include their data in appendices (e-mail 20 July 2001).

Comparative Politics does not have a formal data replication policy, according to Managing Editor Larry Peterson (e-mail 10 July 2001).

Human Rights Quarterly does not have a formal data replication policy (e-mail 27 June 2001).

Journal of Politics does not have a formal replication policy, though editor William Jacoby stated that he hoped to implement one in the near future with the approval of the Editorial Board (e-mail 5 July 2001).

Journal of Theoretical Politics does not have a formal data replication policy. According to editor Keith Dowding, the journal does not normally publish articles that concern the analysis of large data sets. However, Dowding claims the journal will be reviewing its replication policy later this year when the new co-editor, James Johnson, joins him. One of the new features to this journal will be the

availability of articles on the journal's website for those who are required to referee them. This may, as Dowding notes, lead to the availability of data on a similar webpage as well. Dowding also pointed out the contentious nature of replication policies, including the nature of resources to various researchers (less for graduate students, more for professors). In the end, though, the editor claims that "all data should eventually become publicly available, however the time-frame within which this should occur is not obvious," particularly regarding the collection of data for beginning researchers with severe time and resource constraints (e-mail 8 July 2001).

New Left Review There was a bit of confusion regarding our original query. We had no further response from the staff at the journal after clarifying our point. Given the nature of this journal, it seems most likely that it does not have a replication policy.

Political Geography does not have a specific data replication policy. Furthermore, Editor John O'Loughlin claims that no geography journal has a formal policy. Some authors, he states, indicate the location of their data sets in a footnote, though most do not (e-mail 26 June 2001).

Political Theory There was no response to our original query. Given the nature of this journal, it seems most likely that it does *not* have a replication policy.

Politics & Society does not have a formal replication policy. The editor misunderstood our original query (e-mail 26 June 2001).

Public Opinion Quarterly states that submissions not archived at the ICPSR or the Roper Center must contain a definition of the population from which the sample was taken, a description of the selection procedure that allows other researchers an "approximate" replication. In addition, authors must include response rate and calculation details, the dates surveys were conducted, and the precise wording of the questions utilized in the analysis. The equations of the models should be presented, including the numerical values of the parameter estimates, the respective standard errors, and goodness-of-fit statistics for the model, if applicable (<http://www.journals.uchicago.edu/POQ/instruct.html> as accessed 3 July 2001).

Review of International Political Economy We did not receive any further response from this journal. Given the nature of this journal, it seems most likely that it does *not* have a replication policy.

Barriers to Replication in Systematic Empirical Research on World Politics

In spite of decades of controversy on this issue, there is still some obvious inconsistency among those scholars dedicated to the application of systematic empirical methods to the study of international relations regarding the role of "objectivity" in such endeavors. The Peace Science Society (International) declares in its journal *Conflict Management and Peace Science*, for example, that the "Society operates as an objective, scientific organization without political, social, financial or nationalistic bias."⁴ It is probably statements of this sort that lead skeptical observers such as Ann Tickner (1997:619) to declare "empiricist methodologies ... claim neutrality ..." and that the "turn to science" in research on international politics involves the belief that "facts [are] neutral between theories."

In our view, however, it is futile as well as counterproductive for empirically oriented analysts of international politics to strive for, or claim to be "objective." Most analysts, for example, quite clearly prefer peace to war, or the prevalence in the global population of good health and well-being over disease and starvation. These are biased, and un-neutral preferences.

⁴See, for example, the inside cover of *Conflict Management and Peace Science* (Fall 2000).

However, even though researchers inclined to rely on systematic empirical methods cannot convincingly claim to be neutral or unbiased, the *results* of analyses based on systematic empirical methods certainly should be independent of the individual researcher who produced them to an extent that can be ascertained quite precisely by anyone of sufficient competence with the requisite time and energy. That is, the results *are* independent of the researcher who reports them to a degree that is a function of the extent to which they are produced in a manner that is explained in sufficient detail that even passionate theoretical or ideological opponents of the researcher who reports them will be able to analyze the same data in the same way and reproduce the same results. It is this characteristic of systematic empirical research (and not the superior ability of its practitioners to stay above the fray and be objective) that provides (at least in principle) an additional, desirable, margin of separation and difference in the airing of “mere opinions.”

So, in determining the quality of evidence produced by analysts relying on systematic empirical methods, a crucial criterion focuses on the extent to which the evidence in question is produced in such a way that it can be replicated by analysts other than the original researchers who author papers published in academic journals. For the past three years, one of us has been the editor of one such journal in the subfield of world politics, namely, *International Interactions*. It was with some interest then, upon reading Gary King’s piece on replication in political science journals, that the editor came across his discussion of replication policies of various journals in the field. Having noted that such journals as the *American Journal of Political Science*, *Political Analysis*, the *British Journal of Political Science*, and the *Policy Studies Journal* had recently adopted formal replication policies, King (1995:4) concludes that “*International Interactions*, under the editorship of Harvey Starr, is in the process of doing so.”

The editor acknowledges here with some hesitation that this was the point at which he became aware that *International Interactions* does have, and has had for some time, apparently, a replication policy. The “Notes for Contributors” in recent issues of the journal assert the following:

Authors of data-based articles in this journal should indicate how investigators may obtain their data, as well as other information necessary to replicate the reported results, e.g., in which public archives such information has been deposited, or in the case of multiple authors, which co-author to write (not *International Interactions*). The information to be made available should include items such as the original data, specialized computer programs, extracts of existing data files, code books, and an explanatory file which explains how to reproduce the results found in the original article. Authors should remove information from their data sets that must remain confidential, such as the identity of survey respondents. Investigators always have the right to first publication, and, on rare occasions, may wish to embargo the data (or a portion of it) while they finish publishing from it; these authors must specify a date by which a replication dataset will be publicly available.

So, ultimately, the editor is proud to report that *International Interactions* does have a replication policy, and a rather extensive and detailed one at that.

The existence of such a policy is one thing, to be sure, and the degree to which it has been adhered to quite another. One point of this article is to report the results of a modest investigation into the extent to which articles published in *International Interactions* in the years 1999, 2000, and 2001 have adhered to the replication policy that appears in print. In those years, some 46 articles have been published in *International Interactions*. The results of our first step toward ascertaining the extent to which the results of empirical analyses appearing in those articles have been reproducible are presented in Figure 1.

As Figure 1 indicates, in 12 (or 26%) of those articles, the author or authors indicate that the data relied on for the analyses in the article are available online, usually on a webpage maintained by one of the authors. In 8 (or 17%) of the papers published in *International Interactions* from 1999 through 2001, the author(s) asserted, mostly in footnotes, that the data could be obtained by contacting one of the authors. In 5 (or 11%) of the articles published in *International Interactions* during that time period, the data discussed and analyzed were actually presented in the article, in appendices or tables. Six (or 13%) of the articles in *International Interactions* from 1999 to 2001 contained or relied on no quantitative data. Finally, and at least to some extent, unfortunately, the largest category contains 15 (or 33%) of the articles published in *International Interactions* from 1999 to 2001 in which in our admittedly rather hasty review, we could not find any clear indication where or how the relevant data could be obtained.

We continued our investigation by focusing on those papers published in *International Interactions* in the years from 1999 to 2001 falling into categories 2 and 5 in Figure 1. In the first of those sets of papers the author(s) stated that the data were available to those who contacted them, while in the second set we found it at least somewhat difficult to determine how the data in question might be obtained. We e-mailed each of the authors in question a message asking them to help us in our attempt to find out how reproducible were the results published in *International Interactions* in recent years. In short, we asked these authors to send us copies of the data sets in question, along with any supplementary materials that might be useful to any analyst intending to replicate the results they reported.

Five of the 23 authors we tried to contact by e-mail did not respond. These are probably examples of one barrier to replication. If the paper in question was published two or three years ago, and the data are ostensibly available by contacting the author, the author may well have moved. In most cases, additional effort might have enabled any potential replicator to track down the author in question. Three of those contacted said that their co-author actually had the data. We did not follow up in these cases for lack of time. Twelve of the 23 respondents said that in principle they could supply the data, but they did not for some reason or other. Some were just reluctant to go to the trouble to supply the data for what was, quite rightly in their view, a rather formalistic exercise, or simulation. Others said they would supply the data, but it would take time, for some reason or other. We did not follow up by insisting that these respondents comply with our request. Finally, three of our 23 respondents said they had the data immediately available, and they sent us those data as an e-mail attachment.

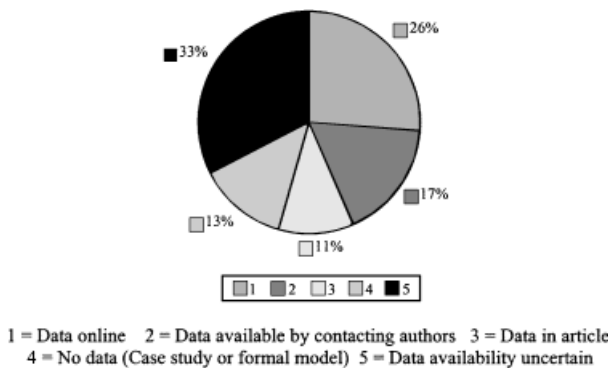


FIGURE 1. Reproducibility of analyses in papers published in *International Interactions*, 1999–2001

This exercise did bring to light different degrees or gradations of “reproducibility,” discussed in some detail by King (1995:451), as well as Herrnson (1995:452). These degrees fall on a continuum from the specific and detailed, at one extreme, to the abstract and formal, at the other. If not only the data set, but also codebooks, and the commands from specific data analysis software, are provided, then an analysis can probably be *duplicated*. Anybody who has performed even relatively simple data analyses will understand just how specific the information will probably have to be if the *exact* same results in the original article are actually going to be produced again by some independent, alternative analyst. So many small changes, recodings, data point modifications, changing of time lags or leads, index renovations, etc., take place in the course of any detailed set of analyses run in the course of producing an article or a book that even the original analyst is likely to forget some of them in a matter of months, if not days. Unless the same data set, complete with codebooks, notes on updating codebooks, and program commands are provided, chances are that attempts at duplication, or replication, will produce different results. And the differences might be substantial.

At the other extreme, analyses are “reproducible” if the explanations or descriptions of the data generation and analysis procedures are sufficiently precise that another analyst can analyze different times and places, depend on different data sources, utilize different software, and produce the same results. If the results are *not* the same under these circumstances, then of course the relevant research community will face the difficult, and perhaps ultimately impossible task of figuring out whether the differences are produced by the fact that the procedures utilized by the analyst attempting to reproduce the original findings were different in some difficult-to-establish ways, or, perhaps, that the original finding does not in fact hold for the different time and place.

Be that as it may, the point here is not that the principle and ideal of “reproducibility” is meaningless. On the contrary, despite the barriers to replication discussed here, reproducibility even in principle represents one of the most valuable characteristics of systematic empirical research results in general, even if in fact, in individual cases, it turns out to be impractical to reproduce given specific results with reasonable investments of time and resources. The optimal reaction to these barriers is not to give up in the face of the problems acknowledged here, but to develop better methods, procedures, customs and disciplinary norms for dealing with them.

One of the more important suggestions made by King (1995:446–447) for improving the replicability of research findings in political science advocates sending the data in some form or other to a “public archive” such as PAVA (the Public Affairs Video Archive at Purdue University), or ICPSR (the Inter-university Consortium for Political and Social Research at the University of Michigan). King rightly notes that authors who notify readers that the data in question are available upon request to themselves have hearts ostensibly in the right places, but “academics are rarely professional archivists,” and “their relatively high degree of mobility also complicates self-distribution, in that affiliations indicated in earlier published articles will not remain accurate.” We encountered problems of this sort in our modest survey of recent authors of articles appearing in *International Interactions*.

However, the senior author of this article is skeptical about the utility of “public archiving” replication data sets at such places as ICPSR. Allow that senior author to acknowledge that these doubts are based on what might fairly be labeled as militant ignorance. He will also acknowledge that he has had no personal contact with the ICPSR for a decade or more. He does have recollections of attempts to take advantage of the services and the data provided by ICPSR. Typically these attempts involved arduous efforts to contact the local representative for ICPSR. At times during the senior author’s career, this “local” representative was located on another

campus several hundred miles away. Once that local representative was contacted, delays of several months were almost invariably promised. On those few occasions when the senior author did actually obtain data from ICPSR, his memory suggests that such acquisitions were invariably followed by immediate discoveries that the data set in question had been updated, and that a better version was available by contacting directly the analyst who had produced or generated that data set. So, over the years, the senior author has come to think of the ICPSR only as a last resort as a source of data, and a fallback option virtually never actually selected. The senior author's admittedly vague and unscientific impression is that his experience has been duplicated among a significant number of IR scholars, so that dependence on the ICPSR among that group of analysts is at a rather low level currently.

A perusal of the ICPSR webpage at <http://www.icpsr.umich.edu/> reveals that at least some of these impressions regarding that archive might be outdated. One of the recent innovations discussed on the webpage, for example, is "ICPSR Direct." According to the webpage:

ICPSR Direct is a new service providing direct access to the ICPSR data holdings for all students, faculty, and staff at ICPSR member institutions. ICPSR Direct represents a quantum leap in data distribution at ICPSR. With ICPSR Direct, any user with an authorized IP address from a member institution may download data from ICPSR, without going through a local campus representative. (<http://www.icpsr.umich.edu/>, 2002)

Such a service would certainly enhance the extent to which ICPSR might serve as an attractive depository for replication data sets. However, we are not convinced that any very general public archive such as the ICPSR is the ideal solution. One reason we reach this conclusion involves a response we received to our e-mail survey of authors of recent articles in *International Interactions*. This particular respondent said he/she could not send us a replication data set, because not all the data were hers to distribute. She explained that she had obtained some (but not all) of the data relied on in the article in question from the ICPSR, and had been told in no uncertain terms that she should not distribute those data to others, that those who inquired should be directed to ICPSR to obtain them.

In situations such as these the depositing of the data set in question at ICPSR constituted a barrier to replication. Even if an interested analyst were to go to the trouble to obtain the data at issue from ICPSR, he or she would then have only a portion of the variables from the data set relied on in the analyses serving as a possible target for replication. Then this person would somehow have to obtain the rest of the data from the author of the original study. Can the author be expected to purge the ICPSR variables from her data set before she sends it to the analyst interested in the replication effort? (Or perhaps, can she take the word of the replicating analyst that he has indeed secured the data from ICPSR, so now can legitimately receive the whole data set, including those data obtained from ICPSR, from the analyst producing the original results?)

Another indication that the ICPSR may not be the (or even an) answer for IR scholars, at any rate, in an effort to move toward higher degrees of reproducibility, is revealed by even a brief review of the "Publications-Related Archive" on the ICPSR website. ICPSR does encourage authors to submit their publication-specific data sets, along with codebooks, and software commands, etc., to ICPSR for deposit at this point on their webpage, just as King (1995) suggests. There are currently some 226 article-related data sets listed there. A very high percentage (90%?) of them are posted by economists. We were able to identify only three posted data sets as clearly focused on international politics. (There were several dealing with international economics.)

In addition, the appearance in recent years of such freely available data management and data-providing sites such as EUGene (Bennett and Stam, 2000), along with such institutionalized sites as those maintained by the Peace Science Society (at <http://pss.la.psu.edu/>) at Pennsylvania State University, and the Integrated Network for Societal Conflict Research at the University of Maryland (at <http://www.bsos.umd.edu/cidcm/inscr/>), where one may obtain such data sets as “Minorities at Risk” and “Polity IV,” seem destined to reinforce the disinclination of IR scholars to become involved with the ICPSR.

It occurred to us while in the process of writing this article that one possible solution to this problem is to have the journals themselves, and/or their editors, maintain websites on which data sets pertaining to articles published in their journal’s pages could be posted. We discovered during that process that this is currently a practice at both the *Journal of Conflict Resolution* and the *Journal of Peace Research*.⁵ Herrnson (1995:454), in an article generally skeptical about anything resembling very systematic or thorough replication policies, argues that “the skills and resources needed to edit a political science journal differ from those used by professional archivists. Just as it is sensible to vest editors with the responsibility for editing journals, it is also sensible to leave professional archivists with the responsibilities associated with archiving data and enforcing archival agreements.” While we would admit that the jury is still out on this issue, we are inclined to believe that journal editors who maintain websites and require authors to post replication data sets on those websites are more likely to raise significantly the levels of reproducibility in the IR subfield in the near future than are those journal editors who encourage authors in the IR subfield to submit replication data sets to the ICPSR. We are pleased to announce here that under new editors (Jacek Kugler and Yi Feng at the Claremont Graduate School), *International Interactions* will establish and maintain its own website with replication data sets. Authors of articles will be required to submit these replication data sets at the point when the transition to the new editors has been completed.

Herrnson (1995) is concerned that a replication policy requiring authors to submit data sets anywhere upon publication of their manuscript risks discouraging researchers from investing time, money, and energy into the creation of data sets. We are certainly in favor of some flexibility in whatever replication policies are adopted in the future to allow analysts to protect their data under some circumstances for some specifically defined period of time. But King (1995:445) points out quite persuasively that the danger to scholars and authors that someone will “steal” their data, and publish astonishing findings before they get the chance, are probably minimal. “The modal number of citations to articles in print is zero; 90.1% of our articles are never cited.”⁶ Again, maybe the experience of IR scholars is different from that of analysts in other subfields; and perhaps our personal experience is idiosyncratic. But we are not aware of even one clear example of an analyst (say, a new Ph.D.) who has had someone get access to their data, and publish results based on that data in such a way as to do clear damage to the young scholar who produced that data. It seems much more likely to us that on occasions when young scholars do have more established scholars publish widely recognized findings based on their data, this helps substantially rather than harms the data producers in question. In short, it seems to us that analysts in the field should have impressed upon them that the risk

⁵While *JPR* adopted this policy in 1998, *JCR* enforced it for the first time in the February 2002 issue. *International Studies Quarterly*, in contrast, requires authors not making their data otherwise available to submit their data sets to the journal. But the editors of that journal also specifically stipulate that “requests for copies of the data must be addressed to the author or authors, and not the offices of ISQ.”

⁶In support of this claim, King cites Hamilton (1991) and Pendlebury (1994).

that they will tightly restrict access to “their” data for months or years only to publish articles or books containing analyses based on those data that will never be cited by anyone, ever, are *much* greater than the risk that some other researcher will mine their data in such a way, and produce publications that will damage their careers.

James Lee Ray
Vanderbilt University

Brandon Valeriano
Vanderbilt University

Replication Policies and Practices at *International Studies Quarterly*

There is only one thing in the world worse than being talked about, and that is not being talked about.

—Oscar Wilde, *The Picture of Dorian Gray*, 1891

Who Needs Replication?

Replication is a scientific ideal, but it also turns out to be good for scholars in practical, even career-oriented ways. Based on a survey of major journals from 1999 to 2001, Gleditsch, Metelits, and Strand (2003) offer striking evidence: an article that makes data available is more than twice as likely to be cited as one that does not. The epigraph from Oscar Wilde certainly applies here; practicing academics hardly could ask for a simpler and more compelling reason to share their data in a user-friendly way—the scientific and pragmatic value of getting “talked about.”

What about the counterargument that sharing data will deprive scholars, especially those who have not as yet made their mark, of the opportunity to analyze the data they have collected first and make the most exciting discoveries? King (1995:445) dismisses the fear of being “scooped” as off base. Some basic points about theory construction would seem to reinforce his argument. Suppose that a data set has “n” variables, and some general principles of organization and current interests in the field, as opposed to an explicit theory, motivated its collection. (This might be expected as something close to the norm for data collection in international relations.) Consider just a few ways in which scholars could vary in their arrangement of the set of “n” variables: exogenous versus endogenous roles; direct versus indirect effects; direction of anticipated substantive impact (i.e., increasing or decreasing); and unidirectional versus bidirectional effects (James, 2002). Moreover, as the “n” gets larger, the preceding possibilities multiply geometrically. Given all of these possibilities, not to mention others, original (or even highly recognized) usage of data is likely to be a product of inspiration rather than the timing of access.

One other issue might be noted: the frequent use of replication in teaching quantitative methods. A typical requirement at the graduate level is to replicate the findings of a published study. Everything, in sum, points toward sharing data meticulously as the best policy for scholars in both scientific and pragmatic terms.

Replication in International Relations and the Social Sciences

Replication is a concept with an ironic place in the scientific study of International Relations. On the one hand, those self-identified with a scientific approach would

seem to speak with one voice about how replication is fundamental to progress. On the other hand, it turns out that the need for replication all too frequently is ignored in practice. King (2003) offers a “replication standard” that seems practical and in touch with general intuition: “Sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party can replicate the results without any additional information from the author.” Yet, as both the present assessment of replication in *ISQ* and studies of other journals (see, e.g., Gleditsch, Metelits, and Strand, 2003; and Ray and Valeriano, 2003) reveal here in this narrative, on average, King’s standard is met neither by the authors of individual articles nor by the editors who are charged with implementing policies.

Reproducible evidence is a virtue frequently extolled but seldom honored among self-consciously rigorous students of international politics—and perhaps that is what should be expected, based on what may be standard practice in the social sciences. As one illustration, the massive *Reader’s Guide to the Social Sciences* (Michie, 2001), with literally hundreds of entries, does not contain one for “replication.” Andrews (2001), who wrote the entry for “evidence” in that same volume, does not discuss replication. While this is hardly a full-fledged survey, the absence of replication as a concept from a compendium such as the *Reader’s Guide* would suggest further that the idea is more of a cherished principle than a way of everyday life among social scientists. Thus the picture would seem to be coming into focus here: it is time for those with a commitment to rigor, regardless of their field of study, to put replication front and center as a priority. Something better is needed.

Hypoprobability and Documentation

Hypoprobability emerges as a key concept that makes the need for thorough documentation very clear. The concept is defined by Cioffi-Revilla (1998:165; see also Cioffi-Revilla and Starr, 1995):

The mathematical reason for hypoprobability is straightforward: probabilities are values defined between 0 and 1, so the multiplication theorem for the probability of a compound event (a mathematical axiom) always requires the product of the n component probabilities, thereby decreasing the value of the probability of the compound event. I call this phenomenon of obtaining a reduced value a *hypoprobability*.

In other words, if a series of events must come true in order for some overall purpose to be achieved, the overall probability is smaller than the lowest probability among the component events.

Hypoprobability clearly factors into replication—achievement of the latter definitely is not a singular event. For example, consider the observations of Ray and Valeriano (2003): “Unless the same data set, complete with codebooks, notes on updating codebooks, and program commands are provided, chances are that attempts at duplication, or replication, will produce different results. And the differences might be substantial.” King (2003) enumerates an even more specific list of needs for proper dissemination of data: “The information deposited should include items such as original data, specialized computer programs, lists of computer program recodes, extracts of existing data files, and an explanatory file that describes what is included and explains how to reproduce the exact numerical results in the published work.” The preceding lists contain four and six components, respectively. Moreover, the components in each instance obviously would entail multiple sub-components if and when the work is carried out as directed. Thus achievement of replication, regardless of which list is used, is a hypoprobable event.

Under such conditions, hypoprobability poses serious problems even if the likelihood of any one component in a series being carried out properly is quite high. Assume, for the moment, that the probability of each step being done correctly is 0.95 and there are either four or six steps, respectively (i.e., two alternative scenarios).⁷ Then the compound probability of overall success in each respective instance is 81.5% and 73.5%. These examples, moreover, are quite friendly ones: a 95% probability of success for each step and just a few to take in order to complete the task of replication. Imagine, then, with dozens of steps and possibly lower probabilities, how hard it will be for scholars, whether experienced or not, to replicate or even duplicate results from minimally complex research designs.⁸ From the standpoint of scientific progress, the inefficiency alone in scholars making isolated efforts to replicate results is a sobering thought. Yet, as Gleditsch and Metelits (2003) point out, “documentation is often thought of as a waste of time.” Again, something better is needed.

Replication at ISQ

What is *ISQ* doing with respect to replication? The following statement appears on the journal’s website and the back cover of each issue:

As part of an ongoing movement toward concerns over community access to data employed in empirical inquiries, and the related possibilities for replication and cumulation, the editors will require authors whose articles are accepted for publication to ensure that their data are fully accessible. If the data in question are not already publicly archived, authors will be required to certify that the data are readily available to others. Requests for copies of the data must be addressed to the author or authors, and not the offices of *ISQ*.

The site also includes a list of data sets used in relatively recent articles. Some entries, at least, appear to be fully in line with the policy. For example, Cioffi-Revilla and Landman (1999) provide explanatory text as well as data. What, then, is the overall picture?

From 1999 through 2001 inclusive, *ISQ* published 84 research articles and notes. Based on the same categories used by Ray and Valeriano (2003) in their analysis of *International Interactions (II)*, *ISQ* ends up with the following distribution: (1) the data set is referred to as being available on line – 10%; (2) the data set is referred to as being available by contacting the authors – 7%; (3) the data set is reproduced in the article itself (e.g., as an appendix) – 7%; (4) no data (e.g., case study or formal model) – 38%; and (5) data availability is uncertain – 38%. Two points stand out in comparing the distribution of articles for *ISQ* and *II*: (a) *II* obviously is a more quantitatively oriented journal, with only 13% of its entries not involving data analysis, as opposed to 38% for *ISQ*, and (b) both journals have a relatively high number of articles for which data availability is uncertain – 33% versus 38%.

Once again, something better is needed—and it now appears to be available. The policy described by Russett (2003) reflects the concerns expressed throughout this symposium and *ISQ* has adopted it. The editors of *ISQ* are committed to the high

⁷In practice, of course, the assumption that component probabilities are identical and independent from each other (i.e., like rolling dice) will not hold. For example, achievement of one task in a series could increase the probability that the next one will be carried out successfully. Alternatively, failure at any stage could make the probability of carrying out all further tasks equal to zero. The more general point to be made is that, as the number of steps increase, the overall probability of success declines, even though the functional form of the connection might not be as simple as the abstract instances provided in the text here.

⁸While anecdotes do not offer compelling proof, two of the students in my graduate seminar last year spent a significant amount of time in efforts to begin projects by obtaining data from authors and replicating results from published articles. The amount of time required to carry out multiple steps in each instance turned out to be substantial—hypoprobability in action.

standards expressed in the policy statement at the end of the symposium and will enforce its implementation.

Conclusions

Produced in consultation with the participants in this symposium, Russett's (2003) prototype for a replication policy represents a way forward. It does not go all the way in meeting Bueno de Mesquita's (2003) call for availability of data and documentation at the time of submission, but perhaps the day will arrive when norms have changed sufficiently among authors and editors to make that practice the standard one. In sum, given the scientific and pragmatic value of replication, its under-utilization in both international relations and the social sciences in general, and the challenges posed to reproducible evidence by the hypoprobable nature of replicating research findings, it is reassuring that something better has been found.

Patrick James⁹

University of Missouri, Columbia

The *Journal of Conflict Resolution's* Policy on Replication

As long-time Editor of a journal publishing quantitative empirical articles, empirical analyst, and producer (alone and with colleagues) of quantitative data I have had a continuing interest in encouraging the ready availability of data and computational routines for replication. So I have been particularly supportive of the movement represented by the articles in this issue of *International Studies Perspectives*. Still, it took that movement to catalyze an explicit requirement about replication for *JCR*. We got serious about the matter with the 2001 volume, requiring authors to state how they would make their data available for replication. But we did not require more than the data alone, nor how material should be made available. Nor did we have any systematic way of ensuring that the requirement was indeed satisfied.

Thanks in large part to discussions with the colleagues represented here in *ISP*, *JCR* started to get even more serious with the first issue for 2002. We created a dedicated *JCR* website for the purpose, and required authors to put their material there. Subsequently we added requirements about posting additional materials to make it fully possible for users to replicate analyses that appear in *JCR*. And we require authors to commit, in print, to doing so as a condition of publication. Although *JCR's* website is meant to be less ephemeral than some, for long-term availability we encourage, though do not require, authors to use other websites as well. We think the additional burden this will impose on each author will be repaid in a much-enhanced process of reciprocation as well as more effective scientific replication. The statement on minimum replication standards at the end of this symposium represents the position that has been in effect at *JCR* for the past year.

So far, all authors have agreed to include the required statement and to post the data. They have indeed provided us with electronic files to post on *JCR's* website (www.yale.edu/unsy/jcr/jcrdata.htm) concurrently with publication. No authors of accepted articles have requested a delay in posting, although in one case an author of a manuscript still under review requested a six-month delay after publication for an especially labor-intensive data set. We judged it a reasonable request, assuming of course that the paper is ultimately accepted. But normally we expect to hold to the requirement of concurrency.

⁹I am grateful to Kim Saak for research assistance. The views expressed in this article are those of the author and should not be taken as a reflection of an official position from the International Studies Association, which publishes *International Studies Quarterly*.

It is true that the specifics of what must be posted are not tight. There may be differences in degree of compliance, pushing us ultimately toward a more standardized requirement. It is also true that, once an article is set in proof, we lose any immediate leverage for ensuring compliance; thus it's possible that someone sometime will take the pledge but not deliver the goods. However, we think such occurrences, if they happen at all, will be rare. We do not wish to slow publication by requiring the material to be submitted to us earlier, for example as a pre-condition of acceptance. Most of our authors seem to accept in good spirit the policy and the reasons behind it. And anyone who fails to fulfill the pledge satisfactorily risks being embarrassed professionally, not to mention establishing a bad record *re* future publication in this journal, or indeed in other journals now that several have already adopted the same policy—and we hope many more will do so.

Bruce Russett
Yale University

Posting Your Data: Will You Be Scooped or Will You Be Famous?

The Replication Debate

In a September 1995 symposium in *PS*, Gary King made a call for replication in the social sciences. He encountered considerable opposition (Gleditsch and Metelits, 2003). The major objection to the proposed replication standard was that it was likely to prove harmful to scholars who have invested a great deal of energy in data collection. Several scholars argued that a replication policy would prevent scientists (particularly younger ones) from making full use of their own data before making them available to others. Herrnson (1995) claimed that the example of researchers who collected data and were subsequently denied its maximum use would deter scholars from investing the resources or shouldering the risks associated with data collection. Forcing individuals to archive their data sets, he argued, would not increase readership. *Original* research questions would increase interest in the field and this would require researchers to collect original data.

In stark contrast to this view, proponents of replication have argued that those who release their data will find that others begin to take an interest in their work, and to cite them, and that this will be to their own advantage (King, 1995:446). Although the call for replication was mainly formulated for the good of the overall research community, those favorable to replication tend to argue that the alleged conflict of interest between the interests of the individual researcher and the interests of the scientific community is a straw man. This article is an attempt to test these two contrasting views.

Testing for Fame

In order to test this hypothesis, we collected citation data to all articles in *Journal of Peace Research* for the period 1991(1)–2001(4), a total of 40 whole issues. We included all types of articles (more below) except the short “Book Notes” (usually under 300 words). Altogether 430 articles were coded. All the citations were recorded within two days (27–28 June 2001) on the citation file maintained by Bibsys, the Norwegian university library system.¹⁰ This file includes *Science Citation*

¹⁰This service was discontinued in early 2002. Norwegian users are now referred to the *ISI Web of Science* at <http://isiknowledge.com/>, which is the standard international gateway to the ISI databases. This database can be accessed by anyone who is affiliated with an institution that holds a current subscription to the databases.

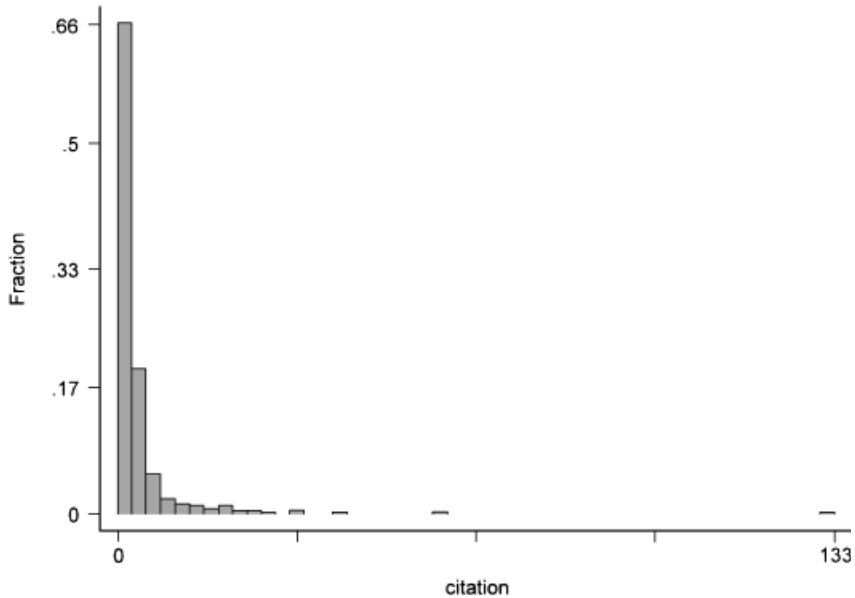


FIGURE 2. Number of Citations per *JPR* Article, 1991–2001 ($n = 430$)

Index Expanded, *Social Sciences Citation Index*, and *Arts & Humanities Citation Index* for the period 1987–2001.

The Dependent Variable

The dependent variable is the number of citations. Figure 2 shows that the distribution of citations is extremely skewed, a common finding in citations studies—whether looking at articles in a journal or researchers in an institute. One-third of the articles were never cited, and two-thirds were cited two times or less.¹¹ The never-cited category was somewhat inflated by including issues close to the citation count. We eventually decided to leave out the 14 articles in the two most recent issues (no. 3 and no. 4, 2001), since none of these had any citations. Thus, in the following tables, $n = 416$.

Independent Variable

To measure replication, we investigated whether the article offered data in any form, either through appendices, URL, or contact addresses. This is a dichotomous variable with a 1 for replication data, 0 otherwise. We used a liberal definition of “data,” but for most of the articles it refers to numerical information in the form of a data set.

Control Variables

It would be unreasonable to expect replication policies alone to determine the fate of individual journal articles. Above all, an article is cited if it is innovative and contributes to the development of the field. This is not easy to capture in itself. However, we have included a number of variables that seem likely to influence the

¹¹King (1995:445) cites studies that show that the modal number of citations to articles in political science is zero: “90.1% of our articles are never cited.” Given that *JPR* is above average in citations (as reported in ISI’s *Journal Citation Reports*), it was to be expected that the share of never-cited articles is lower in our study.

rate at which an article is cited. Some of these are shots in the dark, as will become evident when we look at the results.

Type of article. Since *JPR* is primarily a research journal, we expect regular articles to be cited more than Comments, Review Essays, or the introductory opinion pieces called Focus On until the end of 1997 and more recently Counterpoint. We scored this variable so that a regular article gets a 1 and everything else gets a 0.

Age of the article. The longer an article has been out in print, the greater the probability that someone will have cited it, everything else being equal. Articles published last week are generally not cited by anyone, although with increasing availability on the Web, conference papers frequently get cited and the reference may be amended to refer to a version shortly to appear in print. Normally, however, it would be reasonable to expect an article to have been in print for at least a year before citations begin to appear. Eventually, the novelty will wear off, and the citations will come in more slowly. We therefore include not just age itself, measured as the number of months since the article was published, but also a square term, which we expect to be negative. To avoid multicollinearity, we centered the age variables.

North American author? The citation indexes of the ISI have a bias toward the English language and toward journals published in North America (and particularly in the U.S.). Since all the articles we study here come from the same journal, an English-language journal edited in Scandinavia and published in the United Kingdom, those characteristics do not distinguish here. However, it might be the case that North American authors will more easily be cited, because they have an extensive network of peers who write for North American journals, they have presented their papers at academic conferences in North America, and so forth. This variable has a score of 2 if all the authors are North Americans (including single-authored articles by North Americans), 1 if at least one of several authors is a North American, and 0 if none of the authors are North American.

Political scientist? Like most journals in peace research (and to some extent in international relations) *Journal of Peace Research* describes itself as a cross-disciplinary journal. However, the bulk of the articles (66% in the 2000 volume) were written by political scientists, and this is probably reflected in the readership. We therefore expect that articles from outside might be cited less. This variable also has three values (2, 1, and 0) and is scored like the previous variable.

Male author? We are not truly convinced that gender discrimination is likely here, but it is possible. Since *JPR* records the gender of its authors, we have included this variable. It has three values (2, 1, and 0) and is scored like the previous variable, with 2 for male, 1 for male/female co-authorship, and 0 for only female authors.

Length of article. Longer articles are, *ceteris paribus*, more likely to provide something worth citing. Very short articles are likely to be of less consequence. In some journals (but not in *JPR* in recent years) they might have a special label, like "Research Note." On the other hand, extremely long articles may deter some readers from even starting. We hypothesize that longer articles will be cited more, but that being too long is counterproductive. We therefore include both length and length squared among our control variables.

Degree of formalization of the article. Although methodologically pluralistic, *JPR* has a quantitative orientation. Many readers probably look to the journal to provide

this kind of material. Thus, we expect articles with multivariate analyses of data to be cited more. Even where we have not recorded an article as providing replication data, other authors may in fact have been able to access the data set, and so some of the replication effect should be picked up by the degree-of-formalization variable. We use a five-point scale for the degree of formalization in the article:

- 1: purely verbal argument, though there may be some numbers or percentages, or a list of data without any analysis
- 2: has theoretical figures or tables such as models, flow-charts, etc., and typologies
- 3: has a systematic univariate or bivariate analysis by means of tables, coefficients, or graphs
- 4: contains game matrices using numbers or letters, or figures and equations based on theoretical models
- 5: contains formal multivariate analysis

This coding scheme, which was copied from Gleditsch, Larsen, and Hegre (1994), can hardly pass even as ordinal. We therefore chose to dummy code each of the categories. They are called Formal 1, Formal 2, Formal 3, Formal 4, and Formal 5 in (roughly) increasing order of formalization. In the regression analysis, we include four of these variables with Formal 1 as the reference variable. The lowest degree of formalization is the modal category, with just over half the total number of articles (see the Appendix 3). This category includes most of the non-regular articles (such as review essays).

Co-authorship. It seems reasonable to expect that when several authors cooperate, the result should on average be better than what each of them could have produced single-handedly. Therefore, we include a dummy variable for single-authored (0) vs. co-authored (1) articles.

Appendix 1 provides descriptive information for the variables. There are no missing data.

Model

We analyze the data by means of a negative binomial model. Our dependent variable is an event measure. It does not meet the assumptions¹² of ordinary Poisson regression. Negative binomial regression does not assume independence of events and the distribution is flexible enough to fit our dependent variable (Lindsey, 1995:108). The negative binomial model is interpreted in the same way as a Poisson model, where the expected number of observations is found by multiplying the incidence rate with the exposure time. As our exposure variable is not linear, we control for exposure in the model directly. All observations are assumed to have an exposure equal to 1. In order to compare different independent variables, we compute the incidence rate ratio (IRR), which is the anti-logarithm of the coefficient (*Stata Reference Manual Release*, 1999:29).

Results

Table 2 gives the main results for the multivariate model. Providing replication data does seem to increase the popularity of the article. An author who makes data available is on average cited twice as frequently as an article with no data but otherwise equivalent credentials, including degree of formalization. Empirical articles are significantly more likely to be cited more often, no matter the

¹²This test was performed using the Stata command `poisgof`. See *Stata Reference Manual Release* (1999:425ff) for more information.

TABLE 2. Negative Binominal Model for Citations to *JPR* Articles 1991–2001

	<i>Coef.</i>	<i>IRR</i>	<i>St.e.</i>	<i>p</i>
Replication data	0.83	2.30	0.25	0.001
CONTROL VARIABLES				
Type of article	–.50	0.61	0.23	0.032
Co-authorship	0.12	1.12	0.22	0.598
FORMALIZATION				
<i>Formal 2</i>	0.38	1.45	0.24	0.117
<i>Formal 3</i>	0.70	2.01	0.23	0.003
<i>Formal 4</i>	0.005	1.00	0.34	0.988
<i>Formal 5</i>	0.72	2.05	0.20	0.000
North American authorship				
Mixed, at least one N. American	0.87	2.38	0.37	0.019
<i>All authors North American</i>	0.20	1.23	0.14	0.136
Political scientist/IR scholars				
Mixed, at least one author PolSci/IR	–0.19	0.83	0.34	0.590
<i>All authors from PolSci/IR</i>	0.14	1.15	0.14	0.308
Male authorship				
Mixed gender authorship	0.41	1.51	0.33	0.215
<i>All authors are male</i>	0.23	1.26	0.20	0.254
Age of article				
*Months since publication	–0.03	0.97	0.003	0.000
*Month, squared	–0.0004	1.00	0.00005	0.000
Length of article				
*Length	0.04	1.05	0.04	0.012
*Length squared	–.003	1.00	0.002	0.072
CONSTANT	0.91		0.34	.009
<i>Alpha</i>	1.08		0.110	0.000
Log-Likelihood, constant only model	–954			
Log-Likelihood	–861			
Number of observations (n)	416			
Pseudo-R ²	0.10			

*Centered variable.

sophistication of the method used (Formal 3, 5), and the same is even true of low-level theoretical formalization. Purely theoretical game theory articles (Formal 4) are cited slightly less than the reference category, although the relationship is by no means significant (there are only 24 articles in this category).

Given the cross-Atlantic partnership behind this article, it is inspiring to see that an article with both North American and non-North American authors is cited twice as often as a median publication written without any North Americans. However, this is mainly due to three widely cited articles among the 14 in the mixed category.

Regular articles are, somewhat surprisingly, *less* cited than other publications. This may be an effect of the inclusion of the length of the article, as articles on average are more than twice as long as other items. Articles are significantly *more* frequently cited when we do not use control variables (see Appendix 2).

The length of the publication matters, as predicted, in a curvilinear fashion. The first-order term is positive and strong, while the second-order term is negative and strong, but only significant at the 10% significance level. However, the explanatory power of these two variables is rather small compared to the data and formalization variables. A 20-page publication is cited 15% more frequently than a 14- or 26-page publication, which in turn is 20% more frequently cited than a 10- or 30-page publication.

Other variables perform in the expected way in bivariate comparisons (see Appendix 2) but the effects disappear in the multivariate analysis. In this model, mixed gender articles and articles written by scholars with a political science/international relations background are somewhat more cited, but the relationships are not significant. The same goes for co-authored articles.

Because one article (Jaggers and Gurr, 1995, with 133 citations) was cited a lot more than any other article (no. 2, Oneal et al., 1996, had 59), we reran the analysis with the top-cited article omitted. This only had a limited influence on the coefficients and even less on the significance levels.¹³

We also reran the analysis for just the articles that were published after the replication policy was introduced, that is, from 1998. In this analysis articles with accessible data are on average cited three times more often than other quantitative works in *JPR*. But *n* is much smaller (155) and some of the control variables are no longer significant.

Conclusions

The analysis confirms that making data available seems to improve the number of citations, with a reasonable set of control variables that generally behave in the way that we had anticipated. A weakness of the analysis is that the data replication requirement has only been in place for a few years, so that it still is a bit early to tell whether this pattern will hold. Some articles (including the one most frequently cited) signaled data availability (or published the data in the article) before this became mandatory. Altogether, 75 articles are coded with available data. But there is every expectation that the results for this variable will be strengthened with time as the replication policy has had the opportunity to work for a longer time.

The positive and significant coefficients for empirical articles (formalization levels 3 and 5) point in the same direction. It doesn't seem to matter so much whether the analysis is simple (bivariate) or more complex (multivariate). Empirical articles get cited anyway, whereas purely theoretical formal articles (Formal 4) do not get cited more than the reference category. We compiled a list of the most-frequently-cited articles,¹⁴ and another list of the articles cited more than expected on the basis of the model used in Table 2. The two lists are fairly similar ($r = .51$, with the top-ranking article excluded), which is not surprising given the relatively low overall explanatory power of the model. The top articles in both lists are empirical. Many of these articles will no doubt have inspired attempts to replicate the findings, even if the data were not made available immediately and to everyone. In many cases others will have gained access to the data through correspondence with the author. The clearest example of this is Oneal et al. (1996), the first in a celebrated series of articles on the liberal peace by John Oneal and Bruce Russett. Oneal and Russett have always been generous in sharing their data, and numerous scholars have retested their findings. An article by Barbieri (1996) in the same issue is also frequently cited, but not as frequently as Oneal et al. The data for this article—whose conclusions were quite different—were not at first made available to other scholars. But the differences in the two analyses were fairly clear, so to a large extent they were referred to in the same debate.¹⁵

¹³If our article had been published in *JPR* five years ago, our model suggests that it should have been cited 37 times by now. This is based on an article of 11 printed pages, with cross-Atlantic, multi-gender cooperation with only political scientists, based on available data and multivariate analysis.

¹⁴An earlier analysis of citations to *JPR* articles, covering the first three decades, is found in Gleditsch (1993) and Gleditsch, Larsen, and Hegre (1994).

¹⁵Another striking feature of these lists is that articles dealing with the liberal peace (including the democratic peace) are frequently cited. Out of the top 15 articles cited more than the model would predict, ten deal with the liberal peace and one describes the Polity III data set used to measure democracy in most of these studies. Clearly, this is a reflection of the active debate on the liberal peace in international relations in recent years. But the results would not necessarily be as clear for all other journals in international relations.

We conclude, then, that at least for articles in *Journal of Peace Research*, making replication data available is a policy that seems to serve the author well in terms of citations. Empirical articles are cited more in any case, but in addition there is a net gain in posting the data.

Nils Petter Gleditsch

International Peace Research Institute, and University Oslo (PRIO) and Norwegian University of Science and Technology (NTNU)

Claire Metelits

Northwestern University

Håvard Strand

International Peace Research Institute, Oslo (PRIO) and University of Oslo

Appendix 1. Descriptive Information for the Control Variables

<i>Variable</i>	<i>(n)</i>	<i>Mean</i>	<i>St.dev.</i>	<i>Min</i>	<i>Max</i>
Citations	416	3.48	8.63	0	133
Replication data	416	.17	.38	0	1
Co-authored	416	1.23	.42	1	2
Article	416	.73	.44	0	1
Formalization, level 1	416	.50	.50	0	1
Formalization, level 2	416	.08	.27	0	1
Formalization, level 3	416	.10	.30	0	1
Formalization, level 4	416	.05	.21	0	1
Formalization, level 5	416	.27	.45	0	1
No author North American	416	.53	.49	0	1
Mixed, at least one North American	416	.03	.18	0	1
All authors North American	416	.43	.49	0	1
No author from Pol.Sci./IR	416	.39	.48	0	1
Mixed, at least one from Pol.Sci./IR	416	.05	.21	0	1
All authors from Pol.Sci./IR	416	.57	.50	0	1
All authors female	416	.12	.32	0	1
Mixed gender authorship	416	.09	.28	0	1
All authors male	416	.79	.41	0	1
Months since publication	416	72.8	40.0	2	139
Months since publ., squared	416	6,984	6003.4	4	19,321
Months, centered	416	0	40.0	-70.66	62.34
Months, centered, squared	416	1,601	1417	1.80	4992.46
Length	416	13.73	5.71	1	30
Length squared	416	220.97	161.40	1	900
Length, centered	416	0	5.71	-12.73	16.27
Length, centered, squared	416	32.64	42.59	.071	264.65

n = 416. No missing data.

Appendix 2: Bivariate Coefficients

The following table shows the coefficients for each variable regressed on citations with no other variable but the exposure controls (month centered and month centered squared). For the categorical variables, the frequencies are included.

	<i>Coef.</i>	<i>St. err.</i>	<i>p</i>	<i>(n)</i>
Replication data				416
Data not available (reference category)				344
Data available	.43	.193	.025	72
Formalization, level 1 (reference category)				209
Formalization, level 2	.48	.26	.073	34
Formalization, level 3	1.14	.24	.000	40
Formalization, level 4	.004	.36	.990	19
Formalization, level 5	.91	.17	.000	114
Article				416
Not regular article (reference category)				112
Regular article	.40	.17	.017	304
North American authorship				
No author North American (reference category)				226
Mixed, at least one North American author	1.16	.40	.003	14
All authors North American	.35	.15	.019	176
Political scientist/IR scholars				
No authors from Pol.Sc./IR (reference category)				158
Mixed, at least one author from Pol.Sc./IR	.60	.36	.096	19
All authors from Pol.Sc./IR	.47	.16	.002	239
Gender				
All authors female (reference category)				51
Mixed gender authorship	0.93	.32	.004	36
All authors male	.17	.23	.462	329
Co-authorship				416
Single-authored article (reference category)				321
Co-authored article	.76	.17	.000	95
Age of article				
Months since publication, centered	-.012	.002	.000	416
Months centered, squared	-.0002	.00004	.000	416
Length of article				
Length	.023	.0152	.124	416
Length squared	.0082	.0019	.000	416

Appendix 3. Correlation Matrix for the Independent Variables

	<i>Data</i>	<i>Formal 2</i>	<i>Formal 3</i>	<i>Formal 4</i>	<i>Formal 5</i>	<i>Article</i>	<i>Mixed, at least one N. American</i>
Data	1						
Formal 2	-0.004	1					
Formal 3	-0.21	0.23	1				
Formal 4	0.007	0.12	0.14	1			
Formal 5	-0.33	0.31	0.41	0.22	1		
Article	0.05	-0.20	-0.19	-0.13	-0.37	1	
Mixed authorship, one or more North American	0.21	0.07	0.11	-0.02	-0.04	-0.04	1
All authors N. American	0.07	-0.01	0.06	-0.05	-0.10	-0.10	0.24
One Pol.Sci./IR	0.05	-0.05	0.006	0.03	-0.003	-0.05	0.17
All Pol.Sci./IR	-0.05	0.02	-0.15	-0.01	-0.02	-0.003	-0.08
Mixed gender authors	-0.01	-0.02	-0.19	0.003	0.08	0.02	-0.13
All authors female	-0.06	0.08	0.02	0.01	0.17	-0.12	-0.04
Month, centered	-0.47	0.04	0.09	-0.02	-0.0007	0.19	-0.08
Month sq., centered	-0.31	-0.07	0.18	0.05	0.13	-0.08	0.03
Co-authored	-0.31	-0.10	-0.02	-0.06	-0.21	0.13	-0.33
Length, centered	-0.06	0.07	0.08	0.03	0.07	-0.65	0.02
Length sq., centered	0.01	0.008	0.005	-0.12	-0.12	0.37	0.01
	All North Americans	Mixed, at least one Pol.Sci./IR	All authors Pol.Sci./IR	Mixed gender authorship	All authors female	Months, centered	Months sq., centered
All authors N. American	1						
Mixed, at least one author from Pol.Sci./IR	-0.03	1					
All authors Pol.Sci./IR	-0.25	0.29	1				
Mixed gender authorship	0.07	-0.15	-0.03	1			
All authors female	0.008	-0.06	-0.04	0.57	1		
Month, centered	-0.04	-0.01	0.003	-0.06	-0.01	1	
Month sq., centered	0.002	0.15	0.05	-0.09	-0.05	0.51	1
Co-authored	-0.13	-0.28	0.009	-0.37	-0.11	0.10	-0.009
Length, centered	0.07	-0.04	-0.11	0.04	0.16	-0.28	-0.04
Length sq., centered	0.008	-0.03	-0.02	0.06	-0.03	-0.08	-0.09
	Co-authored	Length, centered	Length sq. centered				
Co-authored	1						
Length, centered	-0.08	1					
Length sq., centered	0.09	-0.18	1				

n = 416

Getting Firm on Replication

International relations scholarship has taken great strides forward during the past several years. New methodological insights have made more rigorous and reliable tests of hypotheses possible. At the same time, the field has also enjoyed explosive growth in cross-sectional and time-series data concerned with almost every facet of international affairs. The quality of research has improved to the point that policymakers are now paying attention to the discoveries resulting from quantitative and formal approaches to international security concerns. This is clearly visible in the policy focus the United States government has developed as a consequence of research on the democratic peace. It is likewise apparent in the use of game theoretic models to facilitate improved decision-making before and during international conflicts and crises and, most recently, in modifying policies for granting foreign assistance. As the impact of international relations scholarship moves gradually from reliance on wisdom to reliance on reproducible scientific knowledge it is important that we, as scholars, and our journals as the conduits for knowledge adopt rigorous standards to ensure that research results can be duplicated and replicated. The hope for greater progress in our understanding of international affairs and for influence over the design of foreign policies demands that we hold ourselves to the highest standards of scientific inquiry.

What Is to Be Done to Advance Duplication?

A healthy trend is emerging among the scientifically oriented journals in international relations to improve the prospects that research results can be duplicated. By “duplicated” I mean that scholars make their data and coding procedures available upon publication so that others can see whether the results are reported with sufficient transparency that they can be reproduced given access to the data and coding rules. This is an important step forward but it is insufficient to promote the greatest possible confidence in research findings. I propose more demanding standards even as I applaud those journal editors who have taken initial steps to promote duplication.

Currently to my knowledge the highest standards applied by international relations journals require that data and detailed coding procedures (as opposed to the more general descriptions often found in manuscripts and published articles) be made available upon publication. In my view, the data and coding procedures ought to be available to those asked to referee a manuscript; that is, upon submission rather than upon publication. Naturally once this step is taken it is essential that referees be limited in their use of the data and coding procedures exclusively for purposes of evaluating a manuscript. Any other use, such as in one’s own research, must require the permission of the author of the manuscript under review at least until after the paper and data are published. With this caveat in mind it seems to me that the fundamental function of duplication is compromised if data and exact coding procedures that move from raw data to constructed variables are not made available prior to publication.

While referees cannot and should not be required to duplicate results, they should have the opportunity to do so. Giving them that opportunity serves several useful purposes. When an author submits a manuscript the data and coding procedures presumably must already be in a useful format. Otherwise the study could not have been done. By requiring that the author submit access to the data with the manuscript editors can ensure that the author has been careful to make procedures transparent because the author knows that there is a risk that a referee will attempt to duplicate results. Thus, the first benefit of such a requirement is that even if referees (or journal staff) rarely actually duplicate a study, the knowledge

that they have the opportunity to do so will discipline authors to be more careful and more precise in their statement of coding rules and in their assembly of data.

Second, by making the data available upon submission, authors provide referees an opportunity to see whether there are significant errors in the data set or questionable decisions in the coding procedures that might compromise the integrity of the findings. Referees cannot adequately identify or offer constructive solutions to these problems if they do not have the data and coding rules in hand at the time they review a manuscript.

Third, referees gain the opportunity to discover whether there are pathologies in the data or variable construction that compromise the interpretation of statistical results. Without an opportunity to evaluate the data and precise variable-construction procedures a referee can only speculate about possible problems with serial correlation, selection/composition effects, colinearity, heteroscedasticity, or a host of other matters. Under current procedures these problems are more likely to be discovered after a study is published rather than before, thereby depriving the author of the opportunity to correct errors and avoid possible embarrassment later on.

Fourth, authors often face revise-and-resubmit decisions based on misunderstandings by reviewers regarding coding rules or the composition of data. Publication is, therefore, delayed while the author responds to queries that might never have arisen if the referee had access to the data and the coding procedures. When a referee believes there are problems in the data or coding rules at present all that s/he can do is speculate about the problem and about ways to fix it. With the data and coding rules in hand, the reviewer can know whether a particular concern is justified or not, thereby improving the quality of reviews and the timeliness of the decision process.

Replication Beyond Duplication

Duplication is an important step toward achieving reliability in research results, but it is only a beginning. For findings to take on the role of knowledge they must be capable of being replicated using different data sets. One reason the democratic peace literature has had a broad impact is that results have been probed and challenged by a large number of scholars. These scholars have used a variety of different measures of democracy and different data sets on international interactions to ascertain whether the key findings hold across time, space, and variations in coding procedures. The democratic peace has been evaluated in contemporary events, but also in events from the 19th century and even from the period of the Greek city-states. Such replications bolster confidence that research findings are not merely tied to a particular moment in time or to a particular way of defining a concept. Yet replication remains relatively rare in our field. What is more, journal editors and researchers alike maintain an unhealthy bias against replication tests in that negative findings are difficult to publish. This too needs to change. If a hypothesis, model, or theory is viewed as a plausible account of how some aspect of international relations works, then findings that falsify or challenge the account are an important part of the scientific process. Indeed, discovery of such falsifying or challenging evidence through replication is fundamental to the acquisition of reliable knowledge.

Conclusions

International relations scholars have an opportunity to set standards for transparency and rigor that are too rarely met in social science research. In my view we should seize this opportunity by requiring authors to submit their coding rules and access to their data at the same time that they submit manuscripts for review. Doing so can improve the quality of the review process and of the resulting

publications. By following the recommendations outlined here we can make progress in speeding up manuscript reviews by eliminating some revise-and-resubmit decisions, providing better information for referees to make more decisive recommendations to editors. Additionally, we can reduce errors in published research and facilitate replication tests designed to probe the robustness of findings. All of these changes will improve the quality and impact of research and so should be adopted as soon as possible.

Bruce Bueno de Mesquita

New York University and Stanford University

The Future of Replication¹⁶

Since the replication standard was proposed for political science research, more journals have required or encouraged authors to make data available, and more authors have shared their data. The calls for continuing this trend are more persistent than ever, and the agreement among journal editors in this Symposium continues this trend. In this article I offer a vision of a possible future of the replication movement. The plan is to implement this vision via the Virtual Data Center project, which—by automating the process of finding, sharing, archiving, subsetting, converting, analyzing, and distributing data—may greatly facilitate adherence to the replication standard.

In King (1995), I proposed that political scientists try to meet *The Replication Standard*, which holds that

Sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party can replicate the results without any additional information from the author.

Meeting the replication standard does not require any work to be replicated. It only requires the provision of enough information so that it could be replicated in principle. Actual replication will often produce other benefits, but the purpose of the standard is only to provide a way to judge the adequacy of the information provided in a published work or a “replication data set” that accompanies the work.

Without information sufficient to make replication possible in principle, the “chain of evidence” from the empirical world to the article’s conclusions is broken, and so the basis for conclusions about the world remains unjustified, at least without taking the word of the author. Although trust is a good thing in personal relationships, science requires that answers depend on public knowledge, not private understandings about individuals and their personalities. The contribution of a work should be represented in that work, not only in the memory or file cabinet of the person who wrote it originally. What counts for the community of political scientists and for our contribution to knowledge and society is only what is on the public record. If they had been academics, Christopher Columbus would have been given tenure early, while Leif Erikson would have been given an extra year to get his book out.

In King (1995) I also proposed some policies for graduate programs, funding agencies, journals, and book presses to encourage or require scholars to follow the standard. Other contributions in this volume report on current practices and some progress at achieving consensus at developing standards. The empirical result in Gleditsch and Metelits (2003)—showing that articles that make data available have twice the impact, and hence make twice the contribution, as those that do not—is

¹⁶Gary King is the David Florence Professor of Government, Harvard University and Senior Science Advisor, World Health Organization (Center for Basic Research in the Social Sciences, 34 Kirkland Street, Harvard University, Cambridge, MA 02138; <http://GKing.Harvard.Edu>, King@Harvard.Edu, (617) 495-2027). The Virtual Data Center project described herein is joint with Micah Altman and Sidney Verba.

particularly striking. Put differently, an author who makes data available has as much impact as two authors who do not. Since we hire, promote, and pay faculty on the extent of their scholarly contribution, it seems likely that those who make data available to others probably also have better jobs and higher salaries.

Science Requires More Than Being Scientific

The training we receive in various methods during graduate school and afterwards is all about making our individual work more scientific, but for a lot of scholars being individually scientific does not add up to a science. Science also requires community, interaction, a common language, the pursuit of common scholarly goals, communication, and a collection of individuals cooperating and competing to contribute to our common and publicly available knowledge base. The tremendous advances in technology, medicine, basic research in the natural sciences, and to some extent research in the social sciences—and the huge increase in the rate of learning in these areas—have come from intense interactions, from checking each other's work, and from individuals building on the work of others.

The famous cold fusion debacle of some years ago involved two researchers, with good reputations, being apparently scientific to the best of their abilities. Fraud was not an issue. Their efforts obviously failed, but the process worked and the world knew the truth within a few weeks. If the same “finding” had been claimed several hundred years ago before modern science (or in a nonscientific field today), society and these individual researchers would have been led astray and little would have been learned. Similarly, when one of us makes an important and accurate discovery about the social or political world, others cannot build on this discovery unless the original author followed the replication standard.

The almost fantastic increase in the rate of mankind's increase in knowledge since the development of modern science is well known. What we may not all appreciate is how recent this all is. Taking the (very!) long term view, the first version of learning on this planet started about 3.8–3.7 billion years ago when the first protocell developed and evolution began (de Dueve, 1995). Evolution, which requires variation plus (natural) selection, produces a type of knowledge as succeeding generations produced organisms that fit various niches and adapted in various ways. The method of learning works, but it is entirely unplanned, undirected, and very slow.

Somewhere around 5–7 million years ago, the last ancestor (the woodland ape) we have in common with our closest living relatives (the chimpanzee) lived. Given the rate at which they learned and their brains expanded, it took until about 1.5 million years ago to tame fire. Creatures then “learned” by evolution, but sometime back then they also started to learn by demonstration. Even today, chimpanzees learn how to use primitive tools and they pass the knowledge to other chimps through demonstration (Wrangham and Peterson, 1996). Most of the knowledge, however, must be learned anew each generation. Around 40,000 years ago (only 2,000 generations) the evidence seems to suggest that language became anatomically possible and subsequently developed. Whatever we learned could now be passed to larger groups without physical demonstration. The result was the onset of extremely rapid cultural development. The form of learning was Lamarckian—the heredity of acquired characteristics through culture—and, relative to evolution, it was very fast and could be intentional and directed. The invention of written language, which permits the transmission of knowledge even if no one presently alive possesses it, dates to 3100 B.C. (or only 255 generations ago).

In contrast, the invention of modern science—with scholars competing and cooperating in the pursuit of common goals on the public record—dates to only about the last 400 years, around 20 generations. It has not yet spread to all fields of

inquiry (e.g., only about a third of the advice you get in your doctor's office is what medical researchers call "evidence-based"), and yet the progress that has resulted far exceeds any other method of learning yet devised (see Haynes, 2002, and the citations therein).

Science includes numerous individual components, and contributions to the infrastructure and definition of science happen all the time. Research libraries date to only about a century ago. Humanity did not know how to run a controlled and randomized experiment, with which one could reliably learn about the world, until the 1930s. Replication and data sharing are even more recently developed norms, and are still spreading across the fields of inquiry. Data sharing in political science is far ahead of medicine and public health but nowhere near some areas of DNA sequencing, where data must be submitted in fixed formats to the same repository as a condition of publication. It is interaction among scholars pursuing common goals that makes science work, and meeting the replication standard clearly advances science.

A Future for Learning from Numerical Data: The Virtual Data Center

Contributing data along with scholarly articles and analyses is obviously in the interest of individual scholars. Much evidence of this point is provided in the other articles in this symposium (and King, 1995). Yet, convincing individuals to do things that are in their long-term interest is often difficult (e.g., think of exercise, recycling, etc.). Helping to develop community-wide norms, instituting requirements when appropriate, and acknowledging and rewarding the data contributions of others can help enormously and should undoubtedly be pursued in every venue available.

In this section, I report on a large software project that attacks the same problem from another angle—with the goal of making it easier for researchers, journal editors, and others to meet the replication standard and improve the product that the discipline of political science provides. Roughly speaking, our project is a hierarchical, structured, and legal version of Napster, but for social science data. The spirit in which we offer this is to make our research lives easier and simultaneously to contribute to the replication movement. In similar fashion, we would all agree to recycle daily if our newspapers and bottles could be separated from our garbage automatically at the landfill without our participation.

The Present and Near Future

The Virtual Data Center, or VDC, is a software project being created by Micah Altman, Sidney Verba, and me (for details see <http://TheData.org> and Altman et al., 2001a, 2001b). It is a project of the Harvard-MIT Data Center and the Harvard University Libraries and is supported by a number of federal grants through the digital library initiative and the political science program at NSF (funding agencies include NSF, NASA, DARPA, NLM, LoC, NEH, and FBI) and part of a separate grant at NIA. I explain what is, what is under way, and what is planned, sequentially.

An operational version of a portion of the VDC is now running at Harvard and MIT and selected other test sites. With this system, faculty, staff, and students knowing only how to point and click can

- search (using a highly focused Google-like system) for data available locally or at various international archives (such as the ICPSR, Roper, NCHS, etc.).
- read abstracts of data sets.
- choose a data set, whether it is local or far away.
- peruse documentation.

- subset the data set by choosing variables and rows (e.g., only women 18–24 who voted for Clinton in 1996).
- perform basic statistical analyses (such as descriptive statistics, cross-tabulations, regressions, etc.).
- download code to replicate the analysis automatically.
- convert the data set format to any statistical package, database program, spread-sheet, etc.
- download the converted and subsetted data for further analyses.
- ensure that the user is authenticated and authorized to have access to the data set chosen.

We can guarantee that VDC software will always be available since it is legally and permanently “open source,” which means that if our group vanishes the software will be on the public record and anyone else will have the right to modify it as they see fit. If they make any improvements to it and offer it for distribution, the license requires them to also provide the source code without charge.

Although many of us grew up specializing in how to extract data from tapes in arcane formats like binary coded decimal or EPCDIC, researchers now call to ask whether they can put their data in our system so they can then take it out in a more convenient format. The usage of the Harvard-MIT Data Center by faculty and students, and for research and in classes, has increased exponentially since its introduction here.

That’s the present. The version to be released shortly will add some critical features. Most important, the same software will be installable anywhere. We have been working closely with the ICPSR and the U.S. Bureau of Labor Statistics, and have had expressions of interest from many universities, governments, and international organizations. When version 1 of the software is released, we expect that many VDC nodes will be installed. This version will make it possible for a user at one site to search across local data sets at their university, data sets at any national archives to which they have access, and data sets at other local VDC nodes. If it works as planned, the system will spread to an ever-increasing user base, and encompass a larger and larger fraction of available social science data. With every additional VDC node, the system will become more useful to everyone.

Others will be able to contribute modules (or add the modules themselves) that snap into the VDC infrastructure and perform specialized tasks, so that many people will no longer need to re-create the wheel at many different sites. Data librarians will be able to build virtual curator’s pages that provide hierarchical organizations and guides for groups of studies in specialized areas, and curators at one site can share the knowledge with others around the world.

As important from the perspective of replication, depositing data will be much easier. We plan to put no hurdles in the way of authors willing to share their data, but the more information (“metadata”) they provide about their data, the more services the VDC will be able to provide users of those data. Perhaps at first data depositors will only drop in zip files with ascii documentation, but we think that when the advantages of all the services of the VDC become more widely known, authors will want to provide more metadata too.

Finally, each data set will have a permanent name associated with it. The name will look like a URL and work in a browser, but it will persist and so will work indefinitely even if the original link vanishes. With this system, checking an author’s claim that data have been deposited will be equivalent to giving a citation to the data. Most journals and copy editors are obsessive about the precise formats of citations to printed matter; this system will make reliable citations to data possible as well.

The Vision: Replication and Deep Citation

The first or second major release of the VDC will also have something like a digital signature associated with each data set and linked to its permanent URL-like name. A digital signature is one number that can be easily calculated from the data set and summarizes all the information in the data set. If any number in the data set changes, then this one number would change too. (An intuitive but bad version of a digital signature, because it is easy to defeat, would be to add up all the numbers in the database.) The advantage of digital signatures is that future researchers could make one easy calculation and then determine immediately and essentially for certain whether they possess the identical version of the data used by the author of the original article. Then we would also be able to verify who created and provided the data. This addition should eliminate an enormous amount of wasted time in building on existing research: a researcher would merely copy down the official name of the data set, type it into a browser, and download the identical data used in the article or book.

Once the names become more widely used, automatic forward citation will be possible. That is, if a researcher comes upon a data set, he or she can instantly find all articles that have subsequently been written using that data set. Ultimately, our goal is “deep citation”: formal ways of citing or referring to particular variables, transformations, or individual cells in the data set. This way, if an author says he regressed income on education, some future researcher will not have to waste a day determining which of the twelve education-related variables were used, and which of the infinite array of transformations that could have been applied were in fact applied. This kind of citation must also be independent of the storage medium, so that if the data set is transferred from SPSS to Stata, the user should still be able to make the right connection transparently and automatically.

Over the years, I have written sample editorial policies for journals that try to help authors meet the replication standard while also trying to accommodate the critics of such policies. The VDC will change these policies considerably. For example, here is one policy that could now be (and in fact already has been) implemented, without the VDC:

Authors of quantitative articles in this journal [or books at this press, or dissertations in this department] must address the issue of data availability and replication in their first footnote. Authors are ordinarily expected to indicate in this footnote which public archive they will deposit the information necessary to replicate their numerical results, and the date when it will be submitted. (The information deposited should include items such as original data, specialized computer programs, lists of computer program recodes, extracts of existing data files, and an explanatory file that describes what is included and explains how to reproduce the exact numerical results in the published work. Authors may find the “Publication-Related Archive” of the Inter-university Consortium for Political and Social Research a convenient place to deposit their data.) Statements explaining the inappropriateness of sharing data for a specific work (or of the necessity for indeterminate periods of embargo of the data or portions of it) may fulfill the requirement. Peer reviewers will be asked to assess the footnote as part of the general evaluative process, and to advise the editor accordingly. Authors of works relying upon qualitative data should submit a comparable footnote that would facilitate replication where feasible. As always, authors are advised to remove information from their data sets that must remain confidential, such as the names of survey respondents.

When the VDC is operational, the policy could be much simpler:

Authors of articles in this journal [or books at this press, or dissertations in this department] that use quantitative data must cite the data by including the Virtual Data Center name for their data and any accompanying replication information.

See your local data center, the ICPSR, the U.S. Census Bureau, or <http://TheData.org> for details of how to do this.

That is, authors merely cite the data in the correct format, just as they are required to cite printed books and articles in the correct format, and the VDC will then automatically take care of the rest.

Concluding Remarks

A warning: beware of vaporware. Although a version of the VDC is now operational at Harvard and MIT, the source code is open and on the Web, and many of the features I have discussed here are already implemented at least in part, much of the rest is presently just promises. Some aspects, such as deep citation, require considerable additional research on our part to address several important unsolved problems. I have enormous confidence in our team of researchers and programmers, but we will have to see.

Whatever the future progress of the VDC, the replication movement will continue. Changing norms and practices is sometimes slow going, but the benefits to individual scholars, to political science as a discipline, and to society at large should keep the movement on its path.

Gary King
Harvard University

Editors' Joint Statement: Minimum Replication Standards for International Relations Journals

Authors of quantitative empirical articles must make their data available for replication purposes. A statement of how that is done should appear in the first footnote of the article. Required material would include all data, specialized computer programs, program recodes, and an explanatory file describing what is included and how to reproduce the published results. This material must be posted by the month of publication, except when, with agreement of the Editor, the deadline is extended to accommodate special need of an author to employ the data for subsequent publications. Information that must remain confidential—such as that which would identify survey respondents—should be removed. All files should be sent electronically to the Managing Editor for posting on a website maintained by the journal for the purpose. In addition, authors may send the material to www.icpsr.umich.edu, and any other sites they wish to use.

We urge other editors to join us in enforcing these minimum guidelines.

Nils Petter Gleditsch, Editor of *Journal of Peace Research*
Patrick James, Co-editor of *International Studies Quarterly*
James Lee Ray, Editor of *International Interactions*
Bruce Russett, Editor of *Journal of Conflict Resolution*

References

- ALTMAN, M., L. ANDREEV, M. DIGGORY, G. KING, D. L. KISKIS, E. KOLSTER, M. KROT, AND S. VERBA (2001a) "A Digital Library for the Dissemination and Replication of Quantitative Social Science Research: The Virtual Data Center." *Social Science Computer Review* 19:458–470. Reprint at <http://gking.harvard.edu/files/abs/vdcwhitepaper-abs.shtml>

- ALTMAN, M., L. ANDREEV, M. DIGGORY, G. KING, D. L. KISKIS, E. KOLSTER, M. KROT, AND S. VERBA (2001b) "An Overview of the Virtual Data Center Project and Software." JCDL '01: First Joint Conference on Digital Libraries. ACM. Reprint <http://gking.harvard.edu/files/abs/jcdl01-abs.shtml>.
- ANDREWS, N. (2001) "Evidence." In *Reader's Guide to the Social Sciences*, vol. 1, edited by J. Michie. London and Chicago: Fitzroy Dearborn.
- BARBIERI, K. (1996) "Economic Interdependence: A Path to Peace or a Source of Interstate Conflict?" *Journal of Peace Research* 33(1):29–49.
- BENNETT, D. S., AND A. STAM (2000) "EUGene: A Conceptual Manual." *International Interactions* 26(2):179–204.
- BUENO DE MESQUITA, B. (2003) "Getting Firm on Replication." *International Studies Perspectives* 4(1): 98–100.
- CIOFFI-REVILLA, C. (1998) *Politics and Uncertainty: Theory, Models and Applications*. Cambridge: Cambridge University Press.
- CIOFFI-REVILLA, C., AND H. STARR (1995) "Opportunity, Willingness, and Political Uncertainty: Theoretical Foundations of Politics." *Journal of Theoretical Politics* 7(4):447–476.
- CIOFFI-REVILLA, C., AND T. LANDMAN (1999) "Evolution of Maya Politics in the Ancient Mesoamerican System." *International Studies Quarterly* 43(4):559–598.
- DE DUEVE, C. (1995) *Vital Dust*. New York: Basic Books.
- FOWLER, L. L. (1995) "Replication as Regulation." *PS. Political Science and Politics* 28(3):478–481.
- GIBSON, J. L. (1995) "Cautious Reflections on a Data-Archiving Policy for Political Science." *PS. Political Science and Politics* 28(3):473–476.
- GLEDITSCH, N. P. (1993) "The Most Cited Articles in *JPR*." *Journal of Peace Research* 30(4):445–449.
- GLEDITSCH, N. P., AND C. METELITS (2003) "Replication in International Relations Journals: Policies and Practices." *International Studies Perspectives* 4(1):72–79.
- GLEDITSCH, N. P., C. METELITS, AND H. STRAND (2003) "Posting Your Data: Will You Be Scooped or Will You Be Famous?" *International Studies Perspectives* 4(1):89–97.
- GLEDITSCH, N. P., T. V. LARSEN, AND H. HEGRE (1994) *Citations to Articles in JPR and by Johan Galtung*. Oslo: Journal of Peace Research.
- HAMILTON, D. P. (1991) Research Papers: Who's Uncited Now? *Science* 4 (Jan.):25.
- HAYNES, R. B. (2002) "What Kind of Evidence Is It That Evidence-Based Medicine Advocates Want Health Care Providers and Consumers to Pay Attention To?," *BMC Health Services Research* 2. <http://www.biomedcentral.com/1472-6963/2/3>.
- HERRNSON, P. S. (1995) "Replication, Verification, Secondary Analysis, and Data Collection in Political Science." *PS. Political Science and Politics* 28(3):452–455.
- INTEGRATED NETWORK FOR SOCIETAL CONFLICT RESEARCH, UNIVERSITY OF MARYLAND (2002) <http://www.bsos.umd.edu/cidcm/inscr/>.
- INTER-UNIVERSITY CONSORTIUM OF POLITICAL AND SOCIAL RESEARCH, UNIVERSITY OF MICHIGAN (2002) <http://www.icpsr.umich.edu/>.
- JAGGERS, K., AND T. R. GURR (1995) "Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32(4):469–482.
- JAMES, P. (2002) *International Relations and Scientific Progress: Structural Realism Reconsidered*. Columbus: Ohio State University Press.
- JAMES, P. (2003) "Replication Policies and Practices at *International Studies Quarterly*." *International Studies Perspectives* 4(1):85–88.
- KING, G. (1995) "Replication, Replication." *PS. Political Science and Politics* 28(3):443–499. Reprint at http://gking.harvard.edu/files/abs/replication_abs.shtml.
- KING, G. (2003) "The Future of the Replication Movement." *International Studies Perspectives* 4(1): 100–105.
- MEIR, K. J. (1995) "Replication: A View from the Streets." *Political Science and Politics* 28(3):456–459.
- MICHIE, J., ED. (2001) *Reader's Guide to the Social Sciences*, vol. 1. London and Chicago: Fitzroy Dearborn.
- ONEAL, J. R., F. H. ONEAL, Z. MAOZ, AND B. M. RUSSETT (1996) "The Liberal Peace: Interdependence, Democracy, and International Conflict, 1950–85." *Journal of Peace Research* 33(1):11–28.
- PEACE SCIENCE SOCIETY—INTERNATIONAL (2002) <http://pss.la.psu.edu/>.
- PENDLEBURY, D. (1994) (Institute for Scientific Information) Telephone Conversation with Gary King.
- RAY, J. L., AND B. VALERIANO (2003) "Barriers to Replication in Systematic Empirical Research on World Politics." *International Studies Perspectives* 4(1):79–85.
- RUSSETT, B. (2003) "The *Journal of Conflict Resolution's* Policy on Replication." *International Studies Perspective* 4(1):88–89.

- SIEBER, J. E. (1991a) Introduction to *Sharing Social Science Data: Advantages and Challenges*, edited by J. E. SIEBER, pp. 1–18. London: Sage.
- SIEBER, J. E. (1991b) “Social Scientists’ Concerns About Sharing Data.” In *Sharing Social Science Data: Advantages and Challenges*. edited by J E. Sieber, pp. 141–150. London: Sage.
- Stata Reference Manual Release(1999) Version 6(3)*. College Station, TX: Stata Press.
- TICKNER, J. -A. (1997) “You Just Don’t Understand: Troubled Engagements Between Feminists and IR Theorists.” *International Studies Quarterly* **41**(4):611–632.
- WRANGHAM, R., AND D. PETERSON (1996) *Demonic Males*. Boston: Houghton Mifflin.

