Quantitative Intelligence Literature: Are Intelligence Scholars Behind the Curve?
ANDREW MACPHERSON, Assistant Professor of Security Studies
University of New Hampshire
andrew.macpherson@unh.edu

ABSTRACT
Is there a quantitative research trend in intelligence literature? Other academic disciplines such as political science and psychology seem to have increased numbers of quantitative articles in their professional journals. In this study, five years of data from the Journal of Intelligence and National Security, International Security, and the Psychological Bulletin journals are analyzed. Articles are categorized as exploratory, descriptive, or explanatory. Articles in the Psychological Bulletin were 60.5% explanatory in nature compared to 8% for International Security and 1.8% for Intelligence and National Security. Why are psychology scholars publishing explanatory work with substantially greater frequency than international relations or intelligence and security scholars? While no single answer explains the discrepancy, three possible deficiencies are explored: academic and methodological training, lack of common definitions, and paucity of data.

Key Words: national security intelligence research, national security intelligence journals, academic explanatory research, national security intelligence literature

INTRODUCTION
Social science academic literature is increasingly quantitative in nature. For at least the last ten years, scholars publishing in disciplines related to intelligence studies such as psychology (Cousinau, 2010), history (Paul et al., 2016), and political science (Abrahms, 2006) increasingly utilize quantitative approaches to develop evidence to scientific questions. Is the same trend seen in national security intelligence literature? If a quantitative research trend is not seen in intelligence literature, might its absence delegitimize intelligence studies as an academic discipline?

Academics studying national security intelligence often receive their training in social science disciplines such as anthropology, history, political science, psychology, and sociology (Smith, 2013). In these disciplines, journal articles seem to be characterized by an increased use of empirical methods. The largest discipline is psychology (National Center for Education Statistics, 2018). Academics with psychology training have a long record of employing advanced analytical methods. In the 1950s, examples of both experimental psychology and correlational psychology were recognized as “streams of method, thought, and affiliation” (Cronbach, 1957, p. 671).

Publishing creative application of analytical approaches facilitates their dissemination and often encourages other academics to employ the similar approaches in their research. Zachary, an anthropologist, formalized an ethnographic description of the fission process as a mathematical model (Zachary, 1977). He visualized the interpersonal conflict in a karate club using a maximum flow–minimum cut Ford-Fulkerson algorithm to accurately predict which members would side with each of the two feuding administrators. Zachary’s data set is available to students and has been used to test new analytical approaches in network science (Girvan & Newman, 2002).
Wilson notes that comparative political scientists moved away from historical methods at the end of the discipline’s formative years, experienced a behavioral revolution that lasted until the late 1960s, and then underwent a second scientific revolution after 1989, which was characterized by greater empiricism (Wilson, 2017). Currently, scholars across most social science disciplines are encouraged to use empirical approaches in novel ways to generate contributions to their discipline (Hider & Pymm, 2008; Zeiler, 2016).

There are at least eight academic journals that focus on national security intelligence issues (Van Puyvelde & Curtis, 2016). The two oldest journals, Intelligence and National Security and the International Journal of Intelligence and Counterintelligence contain over 2000 articles. The methods employed by these journals’ authors are likely representative of intelligence scholars’ the social science training. As social science scholars’ publications generally have become more quantitative in nature, it is possible to compare across disciplines. For example: How does recent intelligence literature compare with near-peer disciplines regarding the use of quantitative research methodologies?

Silke (2001) notes that a broader range of research approaches may facilitate a research field’s ability to “progress from one level of understanding to the next” eventually gaining the ability to explain and predict phenomena. Similar to studies on terrorism scholarship, few analyses exist on the nature and scope of intelligence studies. For example, Smith (2013) provides an overview of the experience and qualifications of intelligence faculty but no insight as to the nature or scope of their research output. Van Puyvelde and Curtis (2016) surveyed the diversity of the authors of academic intelligence studies articles but not the articles methodologies. Johnson and Shelton (2013) survey research topics and gaps in intelligence literature but do not analyze their results by methodological category. Similar to Silke’s work in terrorism studies, the hypothesis developed for this study are designed to benchmark intelligence studies scholars’ exploratory, descriptive, and explanatory publications (Silke, 2001). The goal of this study is to compare the proportion of quantitative vs. qualitative articles published in academic national security intelligence studies, political science, and psychology journals.

**METHODOLOGY**

The journals of Intelligence and National Security (INS), International Security (IS), and the Psychological Bulletin (PB) were selected for analysis. INS is one of the two primary academic journals that focuses on intelligence issues. IS was selected as well, as many intelligence studies scholars received their training as political scientists, of which international relations is a sub field (Smith, 2013). PB was chosen, as it focuses on empirical studies in psychology.

The analysis included slightly over five years of data from January 2014 to May 2019. Journal data was collected and stored in separate Microsoft Excel spreadsheets. The data collected for each journal included the publication name, year, month, issue, and volume, author(s) names; article title; abstract; Universal Resource Locator (URL); and Digital Object Identifier (DOI). Data was collected in May 2019.

A categorization system was developed to code each item retrieved from the journals. Four categories were used for coding: exploratory, descriptive, explanatory, and not applicable (N/A). The first three categories correspond to a research purpose (Robson, 1993). If the journal article primarily focused on proposing theory and exploring basic issues related to the discipline, it was coded as exploratory. If the article contained a literature review, tested theory with case studies, utilized primarily secondary sources, or used survey data to describe phenomena, it was categorized as descriptive. Articles that presented the results of research utilizing quantitative approaches that statistically identified correlations, employed experimental designs, or meta-analysis, were coded explanatory. The final category, N/A, was used as a
placeholder for administrative and other academic journal materials. Academic journals that contain administrative articles such as letters from the editor, corrections, correspondence, and other journal administrative notes. Book reviews were categorized as N/A.

The coding process was content based. The author read the title and abstract to make a preliminary determination of the possible category for the article. If a determination could not be made from the title and abstract, the full journal article was downloaded and reviewed. The Microsoft Excel spreadsheet for each journal was saved as a comma separated value (.csv) file and then imported into the statistic software package R. The software package R is a scripted statistical tool. A scripted statistical tool was selected to mitigate issues related to computational errors that may occur when employing graphical analytical tools such as Microsoft Excel (Panko, 2008). R was used to generate statistics (frequency counts, percentages, etc.) and conduct statistical tests. The imported data was subsetted to remove observations with the N/A label. Table 1 displays the total number of articles for each journal from January 1, 2014, to May 1, 2019, that were included in this study. INS totaled 274 articles; IS, 101 articles; and PB, 248 articles.

Table 1. Total Journal Articles Collected for Analysis

<table>
<thead>
<tr>
<th>Journal</th>
<th>Total Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence and National Security (INS)</td>
<td>274</td>
</tr>
<tr>
<td>International Security (IS)</td>
<td>101</td>
</tr>
<tr>
<td>Psychological Bulletin (PB)</td>
<td>248</td>
</tr>
</tbody>
</table>

A possible research hypothesis could state that INS will have a similar number of exploratory articles as IS. Both intelligence studies and international relations are fields in which theory development is ongoing. I anticipated that INS will have more descriptive articles than IS, as it is likely that data for explanatory approaches is more readily available to political scientists than intelligence scholars. I reasoned that INS will have fewer explanatory articles than IS for the same reason. The hypothesis forms of my reasoning were:

- H1a: The frequency of exploratory articles in INS is not equal to the frequency of exploratory articles in IS.
- H2a: The frequency of descriptive articles in INS is greater than the frequency of descriptive articles in IS.
- H3a: The frequency of explanatory articles in INS is less than the frequency of explanatory articles in IS.

A possible research hypothesis could state that INS will have a greater number of exploratory articles than PB. There may be less theory development in psychology as it is a more mature discipline than intelligence studies. I anticipated that INS will have more descriptive articles than PB for the same reason. It is likely that psychologists are conducting experiments that are generalizable to larger populations, and
thus I reasoned that INS will have fewer explanatory articles than PB. The hypothesis forms of my reasoning were:

- **H4a**: The frequency of exploratory articles in INS is greater than the frequency of exploratory articles in PB.
- **H5a**: The frequency of descriptive articles in INS is greater than the frequency of descriptive articles in PB.
- **H6a**: The frequency of explanatory articles in INS is less than the frequency of explanatory articles in PB.

To test hypotheses H1a, H2a, H4a, and H5a I used a two-proportion z-test. I employed a $p$ value or probability value method, where the $p$ value from the z-test was used to interpret the test results and determine if the null hypothesis was rejected. A $p$ value corresponded to the probability of observing a sample statistic as extreme as the test statistic. When the $p$ value was less than the chosen significance level, the null hypothesis was rejected. I specified a significance level of alpha = 0.05. The 0.05 alpha referred to the probability of rejecting the null hypothesis when it was actually true. The choice of alpha = 0.05 indicated a 5% risk that I concluded a difference in proportions existed when there was no actual difference. I chose not to employ the Yates continuity correction in my calculations (Hitchcock, 2009). As one cell in Table 2 (Explanatory articles for INS) was less than 5, I employed the Fischer exact test for H3a and H6a. The Fisher Exact probability test is non-parametric approach for comparing proportions when the sample sizes are small (McDonald, 2015). To evaluate trends in exploratory, descriptive, and explanatory I generated scatter plots for each journal and plotted the number of articles of each type published that year. I fit a linear regression model to visualize a notional trend line for article type increases and decreases.

**RESULTS**

The total number and percentage of exploratory, descriptive, and explanatory articles for each journal is presented in Table 2.

<table>
<thead>
<tr>
<th>Article Type</th>
<th>Intelligence and National Security (INS)</th>
<th>International Security (IS)</th>
<th>Psychological Bulletin (PB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory</td>
<td>17 (6.2%)</td>
<td>26 (25.7%)</td>
<td>76 (8.9%)</td>
</tr>
<tr>
<td>Descriptive</td>
<td>252 (92.0%)</td>
<td>67 (66.3%)</td>
<td>22 (30.6%)</td>
</tr>
<tr>
<td>Explanatory</td>
<td>5 (1.8%)</td>
<td>8 (8.0%)</td>
<td>150 (60.5%)</td>
</tr>
<tr>
<td>Totals</td>
<td>274 (100%)</td>
<td>101 (100%)</td>
<td>248 (100%)</td>
</tr>
</tbody>
</table>

- **H1a**: The frequency of exploratory articles in INS is not equal to the frequency of exploratory articles in IS.
The two-proportion z-test for H1a produced a p value of 3.67301e-07. The p value of 3.67301e-07 or .00000036 is less than alpha = 0.05 and thus it is highly unlikely these results would be observed under the null hypothesis. The test suggested that the frequency of exploratory articles in INS is not equal to the frequency of exploratory articles in IS. The percent of exploratory articles in INS (6.2%) is less than the percent of exploratory articles in IS (25.7%). Thus, I developed and tested a new hypothesis as follows:

- H1-1a: The frequency of exploratory articles in INS is less than the frequency of exploratory articles in IS.

The z-test for H1-1a produced a p value of 1.836505e-07, which is less than alpha = 0.05 and thus it is highly unlikely these results would be observed under the null hypothesis. Thus, we may reject the null hypothesis in favor of the alternative hypothesis H1-1a. The test suggested that the frequency of exploratory articles in INS is less than the frequency of exploratory articles in IS.

- H2a: The frequency of descriptive articles in INS is greater than the frequency of descriptive articles in IS.

The z-test for H2a produced a p value of 8.984578e-10, which is less than alpha = 0.05 and thus it is highly unlikely these results would be observed under the null hypothesis. Thus, we may reject the null hypothesis in favor of the alternative hypothesis H2a. The p value suggested that the frequency of descriptive articles in INS is greater than the frequency of descriptive articles in IS.

- H3a: The frequency of explanatory articles in INS is less than the frequency of explanatory articles in IS.

The Fischer exact test for H3a produced a p value of .01079, which is less than alpha = 0.05 and thus it is highly unlikely these results would be observed under the null hypothesis. Thus, we may reject the null hypothesis H30 in favor of the alternative hypothesis H3a. The test suggests that the frequency of explanatory articles in INS is not similar to the frequency of explanatory articles in IS.

- H4a: The frequency of exploratory articles in INS is greater than the frequency of exploratory articles in PB.

The z-test for H4a produced a p value of 1. The p value of 1 is greater than alpha = 0.05 and thus we failed to reject the null hypothesis. The test suggested that the frequency of exploratory articles in INS is less than or equal to the frequency of exploratory articles in PB.

- H5a: The frequency of descriptive articles in INS is greater than the frequency of descriptive articles in PB.

The z-test for H5a produced a p value of 5.886069e-80. The p value of 5.886069e-80 is less than alpha = 0.05 and thus we rejected the null hypothesis. The test suggested that the frequency of descriptive articles in INS is greater than the number of descriptive articles in PB.

- H6a: The frequency of explanatory articles in INS is less than the frequency of explanatory articles in PB.

The Fischer exact test for H6 produced a p value of 2.2e-16. The p value of 2.2e-16 is less than alpha = 0.05 and thus it is highly unlikely these results would be observed under the null hypothesis. Thus, we may reject the null hypothesis in favor of the alternative hypothesis H6a, as there is a statistical difference in the proportions. The test suggested that the frequency of explanatory articles in INS is not similar to the number of explanatory articles in PB.
Figure 1 provides a visual representation of the number of exploratory, descriptive, and explanatory journal articles produced annually by INS, IS, and PB during the period of study. The x-axis represents the year of publication and the y-axis is a count of the number of each category of article. A linear regression model was fitted to each category of journal article to approximate a trend line. The bands around the regression lines in the plots represent the range in which the true regression line lies at the 95% confidence level.

DISCUSSION

An analysis of the frequency of exploratory, descriptive, and explanatory academic journal articles in INS, IS, and PB from January 2014 through May 2019 provides data on the state of each discipline. IS contained the highest percent (25.7%) of exploratory articles, indicating that theoretical debate is ongoing in the field. Substantially lower percentages were found in INS (6.2%) and the PB (8.9%). In other disciplines, such as terrorism studies, academics often decry exploratory work as it may encourage “a research community to remain active indefinitely without ever producing meaningful explanatory results” (Silke, 2001, p. 2). Exploratory research is a necessary step in the development of any discipline. One area that warrants future research is the differences in the nature of the exploratory articles in INS and PB. For instance, historical examples of exploratory research in psychology that lead to new descriptive and explanatory studies could be used to stimulate similar advances in intelligence studies. A published comparative analysis may reveal a more nuanced understanding of each disciplines’ research agenda.

Intelligence scholars published in INS employed descriptive approaches much more than contributors to IS or the PB. Descriptive works, such as those employing a literature review, testing theory with case studies, utilizing primarily secondary sources, or using a survey methodology comprised 92% of the articles in INS. Descriptive articles utilizing case studies, survey research, and other methodologies were far less frequent in IS (66.3%) and PB (30.6%). Many of the articles in INS employed a historical
analytical approach where declassified documents and books were fused with insights that address issues in the field.

As an academic discipline matures, it usually develops a corpus of research that forms a matrix of knowledge. Valid and reliable research studies over long periods of time may allow the boundaries of knowledge to advance. Of the three disciplines examined for this paper, psychology appears to be the most advanced as a field of study. Articles in PB were 60.5% explanatory in nature compared to 8% for IS and 1.8% for INS. Many of the studies in PB are meta-analysis. A meta-analysis is an analytical approach where the results from many studies are collected and the results analyzed to generate additional knowledge that may advance the field (Chan & Arvey, 2012).

Why are psychology scholars able to research and publish explanatory work with substantially greater frequency than international relations or intelligence and security scholars? While no single answer explains the discrepancy, I offer three possible deficiencies: the state of academic and methodological training, a lack of common definitions, and the availability of data.

As to the state of academic and methodological training, intelligence studies, as an academic discipline, is relatively new compared to political science or psychology. Some academic and intelligence scholars question if intelligence studies is a discipline at all. Lowenthal noted that the “greatest challenge is gaining acceptance in academia for intelligence as a legitimate field of study and knowledge, distinct from a course or two in the Poli Sci department” (Johnson & Shelton, 2013, p. 116). This study focused on the academics who peruse a research agenda that includes publishing in security related journals. A survey of individuals teaching intelligence-related classes revealed that 48% held social science degrees in political science, international relations, or national security studies (Smith, 2013). While political science methodologies are more quantitative than in the past, they are not comparable to bio statistics or psychology programs that require students show advanced methodological skills to earn a PhD. Some surveys show that political science scholars are increasingly publishing quantitative research (Franzese, 2008). There is a trend toward integrating advanced data science methodologies in academic programs. For example, Massachusetts Institute of Technology offers an interdisciplinary PhD in political science and statistics (MIT, 20019). Graduates with advanced academic degrees that include statistical methodologies who are interested in intelligence studies will likely publish innovative studies that have explanatory power in the future.

Mature academic disciplines establish common paradigms and definitions to support scientific study (Kuhn, 1996). Lack of common definitions may inhibit the development of descriptive studies that may lead to explanatory research (Robson, 1993; Selke, 2001). In the academic field of intelligence studies, many definitions of intelligence and related activities have been proposed and debated. Johnson and Shelton’s 2013 survey notes that defining intelligence is one of the principal scholarly debates advancing knowledge in the domain. The 2006 RAND practitioner/academic workshop titled “Toward a Theory of Intelligence” found no consensus on a definition of intelligence (Treverton et al., 2006).

Individuals employing the scientific method seek to develop relevant and accurate statements that explain the situation of concern or causal relationships of interest. Social scientists using the scientific method develop operationalized definitions (King et al., 1996). Operationalization is a process for assigning rules so a defined phenomenon or object may be measured, or a hypothesis of causal relationships tested. Operationalized definitions facilitate measurement and objectivity (King et al., 1996). A discrete observer should be able to make the same observation or measurement found in a previous study under similar conditions. Beyond technical definitions used for specific scientific activities, broader definitions may be used to generally classify phenomenon or the objects of study. Authoritative literature such as the
Diagnostic and Statistical Manual of Mental Disorders developed by the American Psychiatric Association provides internationally recognized definitions and classifications that aid practitioners and researchers alike. No similar volume exists for intelligence scholars or practitioners. Yet the lack of commonly agreed upon definitions, language, and criteria for intelligence phenomena should not hold back intelligence researchers. In the field of terrorism research, many informative studies exploring causal linkages exist without requiring common definitions. For example, linkages between polity and terrorism have been explored empirically by Li (2005), Piazza (2008), and Chenoweth (2010). A weakness in this approach is the inability for scholars to perform meta-analysis. Meta-analysis relies on common definitions and operationalization of phenomena so study results may be analyzed.

Common data sources may help what Marrin (2016) describes as intelligence studies “generalized failure to ensure knowledge accumulation and aggregation over time.” In the field of terrorism studies researchers, are developing databases that facilitate research into the nature and causes of terrorism. For example, the National Consortium for the Study of Terrorism and Responses to Terrorism maintains an extensive database on 180,000 terrorism incidents from 1790 through 2017 (GTD, 2019). The definitions used to operationalize terrorism and populate the database may allow meta-analysis of studies that employ these data. It seems logical that the same could be done for some intelligence-related topics. However, Zegart (2016) notes that classification and long wait times for government materials deter many tenure-seeking academics from perusing intelligence research and could deter the development of contemporary intelligence studies databases.

Progress could be made. For example, publicly accessible documents on espionage cases in the United States could be assembled from the Public Access to Court Electronic Records system maintained by the United States Courts–Court Services Office (PACER, 2019). This type of data collection is time consuming but is relatively straightforward, as the documents are open source. A database of espionage prosecutions could provide useful data that academics could use for research. Much more challenging—methodologically and logistically—would be a database of interviews with individuals who have been convicted of espionage. Federal laws and regulations govern research with prisoners, who are classified as a vulnerable population (Protection of Human Subjects, 2000, Subpart C). Appropriate training, experience, organizational approvals at multiple levels, and funding would be required to conduct this research.

How does intelligence studies compare as an academic discipline to more mature fields such as psychology and political science? One way of comparing disciplines is to categorize its scholars’ research outputs and compare their frequency. This paper found that intelligence studies scholars produced more descriptive and fewer explanatory journal articles from January 2014 to May 2019 than in the political science or psychology journals. The implications of this research are twofold. First, it is likely that quantitative methods will continue to be applied in creative ways to intelligence studies issues. Quantitative methodologies are prevalent in many of the social science disciplines where intelligence studies scholars receive their training. Second, descriptive research activities may lead to advances in predictive analysis. For example, the development of common data sources and the definitions used to populate them may lead to explanatory analysis. Academic and methodological training, development of common definitions, and availability of data may increase the frequency of quantitative intelligence literature in coming years.
REFERENCES


