Mapping Quality Judgment in International Relations: Cognitive Dimensions and Sociological Correlates

Fabrício H. Chagas-Bastos and Peter Marcus Kristensen

Research quality assessment is a cornerstone of academic practice, yet the criteria that inform such judgments are often assumed rather than critically examined through empirical research. This article draws on a global survey of international relations (IR) scholars (N = 820) to analyze the cognitive dimensions underlying research quality evaluation and their variation across sociological and epistemological factors. We identify seven distinct quality factors: theoretical significance, logical style and structure, practical significance, methodological rigor, contribution and value for future research, interest and topicality, and challenge to existing knowledge. Our results suggest that, while personal preferences, disciplinary norms, and professional practices —shaped by variables such as gender, nationality, and political orientation—influence evaluations, research quality judgments are ultimately grounded in shared cognitive frameworks. Our study offers robust evidence that quality assessments, though subject to sociological variation, reflect deeper, common cognitive structures across scholarly communities.

Assessments of research quality—and cognate standards of "excellent," "ground-breaking," or "original" research—are fundamental to numerous critical scholarly decisions, such as the acceptance of manuscripts for publication, the inclusion of references in syllabi, decisions about hiring and tenure, the awarding of prizes and honors, or the allocation of research funding, all the way to large national research assessment and international ranking exercises. Despite its centrality in academic life, quality judgment has often been taken for granted and has escaped serious empirical scrutiny.

In this article, we map what scholars define as "quality" scholarship. We offer an empirical approach to studying the factors scholars use to judge research quality, focusing on the cognitive structures underlying these judgments—that is, the mental processes involved in perceiving, interpreting, and evaluating the quality of scholarly work—and their relationship to sociological variables. Our objective is to break down the several components (or factors) underlying scholars' information processing and opinion formation toward quality evaluation of a scholarly product. What are the different dimensions researchers privilege in their quality judgments? Also, do such quality standards vary sociologically, and how? Are there gendered or national differences in what criteria are given most weight? Do scholars of different ranks or those working in different theoretical traditions emphasize different quality factors?

Our research lies at the intersection of behavioral science, the sociology of knowledge, and discussions about the political science profession—as well as between the discipline and the broader scholarly landscape in the social sciences. The main contribution of our study is to deepen our understanding of the markers of research quality by uncovering its underlying structures and examining how they vary with sociological correlates such as gender, nationality, ideology, professional rank, and epistemological preferences. Our findings may prove particularly valuable for scholars, especially early-career researchers and those from marginalized groups, as they navigate the academic landscape. Additionally, those in positions of power might use our insights to reflect on whether their

Corresponding author: Fabrício H. Chagas-Bastos (fchagasbastos@usp.br, Brazil and United States) is a research associate professor in the Graduate Program in Latin American Studies at the University of São Paulo. For the 2024–25 academic year, he is a visiting associate professor at the Frederick S. Pardee Center for the Study of the Longer-Range Future at Boston University.

Peter Marcus Kristensen (pmk@ifs.ku.dk, Denmark) is an associate professor at the University of Copenhagen's Department of Political Science.

doi:10.1017/S1537592724002676
© The Author(s), 2025. Published by Cambridge University Press on behalf of American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

positionality influences their perceptions of quality, pondering whether they align with the broader quality standards we present in our study.

In the following sections, we first review existing approaches to the study of quality judgments in academia. In the subsequent sections we outline our methods, and present and discuss our results. Specifically, we explore in those sections (1) the cognitive dimension of quality—exploring the underlying factors that influence scholars' judgments—and (2) their sociological correlates—examining how these underlying quality factors interact with social, geographical, and intellectual contexts. We conclude by discussing the implications of our findings and suggesting directions for future research.

Studying Quality Judgment

A voluminous literature has devoted attention to the study of the social workings of the international relations (IR) discipline (for reviews, see Gofas, Hamati-Ataya, and Onuf 2018; Hamati-Ataya 2012b). Surveys have mapped out the prevalence of certain theories, methodologies, subfields, and policy orientations (Maliniak, Peterson, and Tierney 2012; Maliniak et al. 2017; 2018). Research on the "institutionalized" discipline has investigated how departments, associations, boards, societies, and funding schemes operate (Grenier and Hagmann 2016; Grenier et al. 2020). Work on the "taught" discipline has examined educational practices, such as syllabi and textbooks (Colgan 2016; Darwich et al. 2021; Ettinger 2020; 2023; Hagmann and Biersteker 2014; Murphy et al. 2023; Phull, Ciflikli, and Meibauer 2019). Scholarship on the "published" and "cited" discipline has been prolific in interrogating the discipline's communicative structures (Hvid, Chagas-Bastos, and Kristensen 2025; Kristensen 2012; 2018; Russett and Arnold 2010; Seabrooke and Young 2017; Sillanpää and Koivula 2010; Soreanu and Hudson 2008) and its national origins and geographical representativeness (Aydinli and Matthews 2000; Chagas-Bastos et al. 2023; Kristensen 2015; Lohaus and Wemheuer-Vogelaar 2021; Wemheuer-Vogelaar, Kristensen, and Lohaus 2022). Research on whether citation practices are biased in terms of gender (Alter et al. 2020; Dion, Sumner, and Mitchell 2018; Mitchell, Lange, and Brus 2013; Maliniak, Powers, and Walter 2013; Østby et al. 2013), the prestige of journals, and authors' institutional affiliation (Goh 2019; Hendrix and Vreede 2019) has been particularly prominent.

The sociology of IR literature has thus examined the biases and inequalities in why scholars are published, taught, cited, and so on, but not the quality judgments that underlie such decisions—not to mention tackling it on a large-scale or cross-national basis.[1] Most of these studies—like other areas of research—have largely assumed that "quality" is the residue left when all sociopolitical determinants and biases (e.g., gender, nationality, race) are peeled off, as if "high quality" would then be immediately recognizable to all. This assumption may partly explain why such studies have shied away from studying quality judgment itself. Another factor is the restricted access researchers face when examining quality judgments, due to the inherent confidentiality and anonymity of academic processes such as peer review and selection committees—where criteria and decision making are often opaque, with only editors and panel members fully aware of which factors are prioritized.

Scholarship in other disciplines has explored various approaches to studying quality judgments in research. Broadly, these studies can be divided into two main streams: one focusing on the outcomes of scientific work and another examining the processes behind quality judgments (e.g., how committees decide which research projects to fund).

In the output-oriented stream, a wealth of research has explored the citation count of peer-reviewed articles as a proxy for quality. These studies, for instance, aim to predict the quality of a paper based on a variety of factors such as the characteristics of journals (e.g., high impact factor) or authors (e.g., gender, institutional affiliation, nationality, seniority, and prominence in the field), as well as those of highly cited papers themselves (e.g., their theoretical contribution in terms of content, or other formal items such as their word length or number of references, the number of authors per article), and instances of international coauthorship (for a review, see Xia, Li, and Li 2023; see also Aksnes 2003; Haslam et al. 2008). Using citation metrics as a proxy for quality assessment is a relatively consistent method for identifying high-quality articles based on quantifiable criteria. The relationship between quality and citation count is, however, dubious at least (see Herrmannova et al. 2018). Certain articles may experience a short-lived surge in citations before fading into obscurity, or they may be frequently cited for their shortcomings rather than their strengths—which does not tell us much about their quality. Gottfredson (1978), for instance, found that experts' assessments of various dimensions of an article's quality were weakly correlated ($r \leq 0.23$) with citation counts. Similarly, Shadish (1989) found that although psychologists' ratings of article quality indicators were significantly predicted by their assessments of 25 out of 27 quality-related criteria, only four of these criteria were shown to predict citations ($r \leq 0.22$). Furthermore, Lee and colleagues (2003) found that prospective judgments of quality in the form of "outstanding paper" awards only weakly predicted eventual impact. In short, bibliometric studies can offer insights into strategies for accruing citations (e.g., Baldi 1998). However, they are neither a reliable nor a direct measure of research quality and, more importantly, fail to uncover the reasoning behind decisions to cite or ignore particular works.

In the process-oriented stream, we find studies that examine how scholars make judgments in practice—for

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

example, prize and funding panels, editorial boards, and review processes. We can distinguish here between two types of quality judgment: *retrospective judgment* focuses on completed work (e.g., the assessment on whether to accept or reject a manuscript for publication, or whether to award tenure) and *prospective judgment* concerns how promising work may be in the future (e.g., reflecting on whether to fund proposals that might eventually lead to groundbreaking research). Guetzkow, Lamont, and Mallard (2004, 191; see also Mallard, Lamont, and Guetzkow 2009) showed that judgments of quality in the social sciences and humanities often rely on interpretively flexible criteria, such as originality and significance, which can mean anything from "using a new approach, method, or data, studying a new topic and doing research in an understudied area, as well as producing new theories and findings" to having a "grasp of the relevant literatures" and displaying "scholarly excellence" (Lamont 2009, 27, table 2.1). The seminal sociological work of Michèle Lamont (2009) showed that judgments of quality in grant review panels are relative (to disciplinary environments, among other things) and more cognitively, socially, and emotionally attuned than scholars might like to admit. In the same vein, studies of the peer review process have compared original submissions, reviewer reports, and final publications to examine what criteria reviewers actually use, and have found that revisions often aim more at theoretical or conceptual reaming, reworking the discussion and literature engaged (i.e., interpretive revisions), rather than revisions of methodology or data (Strang and Siler 2015). Although output-oriented studies demonstrate strong reliability and generalizability, their validity is undermined by a tendency to conflate quality with citation counts. Moreover, these quantitative studies often obscure the process of quality judgment itself. In contrast, process-oriented studies aim to unpack the black box of quality judgment. However, these qualitative studies generally suffer from weaker reliability due to the challenges in accessing judgment processes, often relying on analyses and interpretations of scholars' perspectives as expressed in their publications or through interviews.

A Cognitive Approach to Quality Judgment

Psychologists Robert Sternberg and Tamara Gordeeva (1996) proposed a different approach to quality judgment, exploring the cognitive processes involved in how individuals think and decide about quality. To study the anatomy of research "impact" and uncover what distinguishes influential papers, they drew on extensive research in intelligence (see Sternberg 1997 for a review). The foundation of their approach is the *triarchic theory of intelligence* (Sternberg 1985a; 1985b; Sternberg et al. 1999), grounded in the idea that intelligence encompasses three key cognitive dimensions: *analytical reasoning*, the critical evaluation and logical processing of information; *creativity*, the ability to generate innovative ideas and tackle problems from unconventional angles; and *practical problem solving*, the capacity to apply knowledge effectively in real-world contexts.[2]

Based on the idea that assessments of research are also a product of underlying cognitive dimensions, they presented scholars with a 45-item questionnaire focused on the general properties of a research article, asking them to rank these properties by their importance for making an "impact" on the field—a proxy for quality. As a result, they identified six factors driving influential research: (1) quality of presentation; (2) theoretical significance; (3) practical significance; (4) substantive interest; (5) methodological rigor; and (6) value for future research (see appendix C in the online supplementary file for further details). This six-factor model reflects what scholars might consider when evaluating the potential quality of research, and may intuitively align with what characterizes impactful work. Above all, they provided a sophisticated and comprehensive framework for evaluating the impact and significance of scholarly work, highlighting the nuanced nature of quality judgment in academia. These factors offer a comprehensive, albeit not exhaustive, list of elements that give us insight into what the scientific community might see as the underlying features of the most impactful articles in the field. The merits of this cognitive approach, compared to existing work, lie in its ability to investigate the building blocks of what constitutes quality judgment in a replicable manner.

Building on Sternberg and Gordeeva's work, we study a selection of published IR authors and their underlying features of quality judgment. While their study primarily examined the most salient criteria within "impactful" peer-reviewed journal articles, our approach focuses on what distinguishes *high-quality* ones. Impact is more interpretively flexible than asking directly about quality.

We aim to fill two main research gaps by examining, first, the cognitive structures of quality judgment within IR and, second, how these quality standards vary according to scholars' positions in the discipline. First, individuals (researchers) with different cognitive styles might prioritize different criteria, interpret evidence differently, and arrive at varying conclusions when evaluating research quality.[3] We have limited knowledge about the *underlying quality features* scholars prioritize when evaluating research. Are these criteria tied to the novelty of the theoretical argument, methodological rigor, clarity of ideas, or the potential to initiate new research areas? Could it be policy relevance or perhaps the alignment of the research with existing literature? In short, we need to have a clear understanding of the underlying *cognitive dimensions* of quality judgment in IR.

Second, our purpose is to gain an insight into how cognitive dimensions of quality judgments vary according to sociological correlates. Cognition does not exist in a

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

vacuum. We know even less about whether and how such cognitive quality standards *are embedded sociologically.* How do quality judgments vary in terms of gender, geographical background, ideological beliefs, or professional seniority? Furthermore, how do these quality factors vary along with different theoretical leanings, methodological preferences, and fields of study?

Sternberg and Gordeeva provide the tools for cognitively mapping quality judgment, and our study pushes research forward by examining the sociological correlates of these latent quality factors. A thorough comprehension of the cognitive structures of quality and how they interact with the positionality of researchers enhances our understanding of quality judgment in IR, and also provides a novel take that can be replicated in other branches of social science.

Methods

We gathered our data between December 2022 and January 2023. Ethics clearance for our study was granted by the Research Ethics Committee at the Faculty of Social Sciences at the University of Copenhagen and Horizon Europe's Marie Skłodowska-Curie Actions (grant agreement no. 101032425). Further details on statistics, demographics, and complete question wording are reported in the online supplementary file to save space; for the raw data, see Chagas-Bastos and Kristensen (2025).

Analytical Strategy

We employed a two-step process to identify cognitive factors underlying quality judgments and examine their relationship to sociological variables in the context of IR research. The first step involved identifying latent cognitive quality constructs through exploratory factor analysis (EFA) to determine the key elements scholars associate with quality research. Following this, we utilized the scores of each quality factor in our regression models to evaluate how quality factors and sociological variables are interrelated.

Power Analysis and Sampling

Given the uncertainties about the size of our target population, we adopted a sampling approach based on statistical power calculatios to ensure the robustness of findings overall and at each stage of analysis.[4] Guidelines suggest a ratio of *at least* 10 participants per survey item for the EFA (Bryant and Yarnold 1995; Comrey and Howard 1992; Gorsuch 2015). Given our 49-item questionnaire (see details below), this translates to a minimum of 490 participants. For our regression models, we conducted an a priori sampling calculation using `R` software.[5] Our goal was to achieve high power (0.95) to detect small effect sizes ($\beta = 0.20$) at $\alpha = 0.05$ (two tailed) in multiple linear regressions with nine predictors.[6] The results indicated a minimum required sample size of 133 participants. To account for potential data loss and to ensure robustness in both techniques used, we decided to collect data above the highest minimum threshold.

Recruitment

We also employed an inductive recruitment strategy to cast a broad net and engage as many participants as possible from the global IR scholarly community. To compile a list of potential participants, we gathered publicly available email addresses of authors who had published peer-reviewed journal articles in the "international relations" and "political science" categories[7] of the Web of Science database, covering the period from January 2000 to December 2022.[8] Published authors who had undergone peer review in IR journals were considered the relevant population for our study, as they represent a significant cross-section of the global IR scholarly community. We recognize, however, that many IR scholars worldwide may not have published in Web of Science-listed journals, and some who have may not identify as IR scholars. Furthermore, recruiting solely from this pool would also introduce an Anglo-European bias. To broaden the scope of potential participants, we manually collected email addresses of colleagues in IR and political science departments, as listed on institutional web pages, from four of the largest academic communities in the Global South: Brazil, India, Mexico, and Turkey.

Demographics

A total of 820 individuals from 83 countries participated in our research (32.1% female; 90.2% held a PhD; 85% self-identified as working in an IR subfield). While we do not claim that our sample is representative of the IR discipline, the demographics align with those of comparable studies (see appendix B in the online supplementary file for details). Political ideology distribution in our study, for instance, follows what has been found in previous research (e.g., Rathbun 2012). Similarly, gender distribution aligns with the 2017 (total sample) and 2022 (United States sample) waves of the TRIP survey (Entringer et al. 2023; Maliniak et al. 2017). Our primary focus is on individual-level cognitive differences, rendering concerns regarding the representativeness of the sample less important. Still, our sample size ($N = 820$) allowed for 95% statistical power to detect small effects (two tailed) in ordinary least squares (OLS) models and meets the requirements for a robust EFA.

Measurements

To assess the underlying factors of research quality, we adapted the measurement developed by Sternberg and Gordeeva (1996). We made minor textual modifications

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

to the original questionnaire, merely rephrasing items or removing disciplinary specificities to make them broadly applicable to other areas of science. Additionally, we removed four items that were specific to psychology and replaced them with eight new ones, ensuring the scale is as inclusive as possible. The revised set comprised 49 statements about quality research, answered using a seven-point Likert scale (1 = "not at all important," 4 = "neutral," 7 = "extremely important"). Given that the original text instructing participants about the task asked respondents what they considered necessary to make an "impact in the discipline," we revised it to explicitly focus on quality.

When asked about impact, some respondents might focus on what is needed to accumulate citations, others might consider the criteria for a paper to gain fame, and still others might point to different indicators of "impact in the field." To address this limitation in the original study, we explicitly asked scholars about "quality," aiming to avoid the interpretive flexibility that could lead respondents to base their answers on idiosyncratic understandings of impact.

Procedure

Participants were contacted using an automated mailing system in three waves between December 2022 and January 2023 and were invited to participate in our study. Those who expressed interest in participating in the study were directed to the web-based Qualtrics survey software to read and complete both the plain-language statement and the informed-consent forms. They were explicitly told that their participation was voluntary and anonymous, and that they could opt out at any time without any penalties. They first answered demographic (age, gender, nationality), professional (professional rank, employment location), and epistemological (theoretical paradigm preferences and areas of study) questions.[9] Next, participants completed the questionnaire about research quality and, finally, political orientation. The order of all measurements and items in each questionnaire was randomized. After participants completed and submitted their responses, they were debriefed about the study and informed of their contributions to it. The response completeness rate was 100%.

The main questionnaire in the study asked scholars about their "views on the importance of the parameters below when evaluating the quality of work in International Relations" and to rank what they consider important. We essentially asked participants to adopt the perspective of a reviewer, focusing on the standards they personally deem crucial in assessing quality (*individual* quality standards), rather than the perspective of the reviewee assessing the standards by which their own work or research in general is evaluated during review (*communal* quality standards), as follows:

> This questionnaire seeks your views on what importance the parameters below hold when evaluating the **quality of work** in International Relations—and more broadly Social Sciences. Each statement represents an attribute contributing to high-quality levels of a **journal article** studying world politics. Your task is to rate on a 1 to 7 scale, where 1 indicates that you do not believe that the attribute is of any real importance in determining the quality of an article, and 7 indicates that you believe that the attribute is of extreme importance.

Control Variables

We controlled our models for basic demographics (gender and nationality) and political orientation. The Social and Economic Conservatism Scale (SECS; Everett 2013) was used to measure participants' multiple ideological dimensions. The SECS consists of 12 items ($\alpha = 0.91$) that relate to economic (e.g., "limited government") or sociopolitical (e.g., "religion") issues. Participants rated their attitude toward each issue on a hundred-point scale, where a score of one hundred indicates a greater level of positivity (i.e., high conservatism) and a score of zero indicates a greater level of negativity (i.e., low conservatism).

The Cognitive Dimensions of Quality Judgments in IR

The top 10 ranked quality items are oriented toward writing, presentation, and comprehension, with aspects such as clarity, structure, flow, succinctness, and consistency followed by others centered on the manuscript's significance, its potential contribution to the field, and contextualization within existing literature. In the following items (ranks 11–18) the emphasis shifts to novelty, with quality items such as presenting new ideas, offering new and better explanations, and opening new research avenues. Next, we find items (ranks 19–25) centered on capturing attention, exemplification, clear messages, and broad relevance. The items ranked 26–36 are oriented toward theoretical and conceptual contributions, theory building, and the presentation of new, alternative, or modified concepts and theories. Among the least important items, we find items related to methodological and data contribution as well as practical and policy significance (ranks 37–41), and in the lowest ranks (44–49) ones such as debunking or falsifying existing theories, timeliness, and the generalizability of findings and theories.[10]

The analysis of standard deviations showed a high level of consensus among participants on items related to the logical clarity and coherence of argumentation (items 1, 2, and 4). The focus on items related to presenting and organizing ideas or results seems to reflect a broader trend in the social sciences, where academic writing is expected to adhere to a *scientific* style. Items with the least agreement concern the use of hypothesis testing (item 34) and an unbiased tone (item 31), as well as timeliness (item 48). The high level of disagreement observed on the former two

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

can be taken as an indication that the positivist–postpositivist divide in the discipline transfers into quality judgment.

To identify how these latent quality constructs cluster, we used the same EFA methods (i.e., principal component analysis with varimax rotation; see appendix D in the online supplementary file for further details) as in the original study. Our analysis successfully replicated the original six-factor structure, albeit with minor revisions in their labels, and revealed a novel factor (factor 7, items 39, 35, 38, and 45). Table 1 displays the factors ordered from the highest to the lowest eigenvalue—that is, the strength of the factor in accounting for variance within the correlational data.[11] Taken together, the seven factors in our study explain 54.46% of the variance, compared to 49.76% in the original study.

Table 1
EFA for the Latent Cognitive Quality Constructs

Scale value	Loading	Item
		Factor 1. Theoretical significance (Average scale: 4.99; eigenvalue: 11.91; Cronbach's α = 0.89)
5.00	0.78	Presents a useful new theory or theoretical framework
5.16	0.76	Presents results of major theoretical significance
5.15	0.74	Contains useful implications for theory building
5.14	0.73	Presents a useful new concept or theory
5.18	0.69	Proposes alternatives to an existing theoretical and conceptual framework
4.92	0.61	Contains some surprising results that make sense in some theoretical context
5.05	0.61	Contains useful recommendations for changes or modifications in accepted theoretical constructs
4.57	0.57	Debunks an existing theory or way of thinking
4.70	0.50	Develops a new and useful methodology
		Factor 2. Logical style and structure (Average scale: 5.96; eigenvalue: 4.04; Cronbach's α = 0.86)
6.15	0.74	Has a logical flow and organization of ideas
6.15	0.72	Well written, well structured, and well organized
6.21	0.71	The problem is clearly stated and well conceptualized
5.81	0.70	The writing is succinct and internally consistent
6.19	0.70	Arguments are presented clearly and discussed carefully with tight, logical reasoning
6.08	0.62	Results are presented clearly and discussed carefully with tight, logical reasoning
5.58	0.59	The language meets acceptable academic standards
5.71	0.56	There is adequate reference made to other work in the field
5.73	0.48	Contains useful implications for a scholarly understanding of the field
		Factor 3. Practical significance (Average scale: 5.01; eigenvalue: 2.20; Cronbach's α = 0.79)
4.73	0.72	Contains useful implications for policy and professional practice
4.83	0.68	Results are of major practical significance
4.63	0.54	Is applicable to work in many other research areas
5.25	0.54	Is clearly understandable to a broad cross-section of scholars
5.39	0.48	Includes concrete examples
5.21	0.43	Contains useful recommendations for further research or for changing research methodology
		Factor 4. Methodological rigor (Average scale: 4.82; eigenvalue: 1.87; Cronbach's α = 0.77)
5.02	0.74	Hypotheses are clearly stated and testable
5.09	0.64	Tone is unbiased and impartial
5.04	0.56	Contains interesting generalizations, which are clearly stated, confirmed, and based upon results
3.88	0.47	Presented theory is a general one
5.22	0.46	Considers different alternative interpretations of the data and then arrives at a clear final, unambiguous interpretation
4.66	0.45	Integrates into a new, simpler framework data that have previously required a complex and possibly unwieldy framework

(*Continued*)

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Table 1 (*Continued*)

Scale value	Loading	Item
		Factor 5. Contribution and value for future research (Average scale: 5.54; eigenvalue: 1.25; Cronbach's α = 0.72)
5.44	0.64	Generates new research
5.53	0.63	Opens up a new problem (research question) for investigation
5.51	0.46	Contains useful implications for future research studies
5.69	0.44	Provides a better explanation of existing phenomena
		Factor 6. Interest and topicality (Average scale: 5.24; eigenvalue: 1.08; Cronbach's α = 0.71)
5.39	0.63	Captures reader's interest
5.74	0.60	Topic is interesting and important
4.02	0.52	Appears at the right moment, when people are ready to hear the message
5.33	0.49	Starts and ends strongly, attracting attention and interest from the first paragraph and ending with a clear take-home message
5.69	0.47	Provides new and exciting ideas
		Factor 7. Challenge to existing knowledge (Average scale: 4.49; eigenvalue: 1.05; Cronbach's α = 0.63)
4.34	0.71	Results are intriguing but do not fit any existing theory
4.41	0.54	Provides evidence that fails to support an existing influential theory
4.06	0.47	Presented findings are general ones
5.15	0.45	Critiques existing knowledge

Note: Items were listed only if their factor loading was 0.40 or greater.

The factors are interpreted as follows:

- Factor 1. *Theoretical significance* contains a range of quality items that emphasize innovative ideas or alternatives to established theories, and contains novel results that make new sense of or debunk existing theoretical frameworks. This can be taken as evidence that theoretical significance is still a major factor in IR when its scholars assess the quality of manuscripts, regardless of consistent concerns about the "end" of IR theory (Dunne, Hansen, and Wight 2013; Guzzini 2013; Mearsheimer and Walt 2013; Sylvester 2007; Wæver 2016). The premium on theoretical significance, however, does not automatically imply that the theory advanced in a manuscript needs to be contributing to the grand IR "isms" or the so-called great debates. Rather, as the top items in factor 1 indicate, it is the *novelty* of theories or concepts that is perceived as a central quality criterion. This finding aligns with the argument made regarding the proliferation of new theoretical "turns" in IR (Baele and Bettiza 2021; Heiskanen and Beaumont 2024).[12] Although making a novel theoretical contribution might be a hard task, importing a novel concept or theory from other fields or disciplines is perhaps less so—and it seems to be the path taken by some of the most successful approaches in IR. The integration of external concepts can serve as a catalyst for theoretical innovation within political research.
- Factor 2. *Logical style and structure* contains items related to argumentation and presentation: clear, logical organization and progression of ideas, clear problem statements, succinct and consistent (i.e., unambiguous) writing, and proper academic language and referencing to relevant literature. Writing is not just about grammatical correctness and style, but also a culturally variable conception of what counts as logical flow, eloquence, and well-organized and solid argumentation. A clear writing style is sometimes viewed as a direct indicator of personal intelligence, scientific competence, and "clarity of the mind" (Lamont 2009, 168). On the surface, the importance of a logical style and structure seems uncontroversial; it is not, however, as innocuous as it seems. Upon closer inspection, the items in factor 2 indicate that clarity is specifically associated with a standardized writing style that previous studies in the sociology of IR have labeled as distinctly Anglo-Saxon, one that mimics the argumentative structure and rhetoric found in the natural sciences "with brief, straightforward statements and linear progression of an argument" (Wæver 1998, 694; see also Horn 2017). This style has arguably become hegemonic (at least in the West), at the expense of other forms of academic expression such as, inter alia, the more complex (and precise) German writing style with attached provisos (Wæver 1998), the French style of argument structuring and allusion (Breitenbauch 2013; Lamont 1987), or a more holistic and circular "Chinese thinking style"

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

(Kristensen and Nielsen 2010, 73), so much so that our results showed that IR scholars worldwide designate an Anglo-Saxon writing style as a token of quality.

- Factor 3. *Practical significance* contains items suggesting that quality in a manuscript also hinges on presenting results of practical significance, with useful implications for the academic profession and for policy, as well as ideas that are applicable to many areas (across the discipline and beyond). It suggests that quality research is embedded in the environments in which research takes place, including the problem-solving context and societal influences, and shapes the choice of research topics and design, as well as the potential uses of the findings (Gibbons et al. 1994; Nowotny, Scott, and Gibbons 2003). This factor supports recent research (Hendrix et al. 2023) that has debunked the long-standing narrative about IR as a "cult of irrelevance" lacking serious engagement with policy (Avey and Desch 2014; Avey et al. 2022; Desch 2019). Our findings provide further support that practicality, or perhaps the perception of policy relevance, is also seen as an intrinsic part of research quality.
- Factor 4. *Methodological rigor* refers to items converging on scientific rigor and objectivity. The first items stress clear and testable hypotheses, impartiality, and evidence-based generalizations, while the last two items specifically concern empirical contributions. This factor highlights that quality in a manuscript can also stem from presenting innovative methodologies, novel data, or cutting-edge research techniques, as well as from distilling otherwise complex data into a comprehensive framework. The underlying components in factor 4 resonate with debates over whether IR scholars increasingly favor methodological rigor or political relevance, and "simplistic hypothesis testing" over theory building (e.g., Mearsheimer and Walt 2013; Walt 1999).
- Factor 5. *Contribution and value for future research* points toward the key role of originality and innovation in a manuscript. The synthesis of items in this factor reflects the value placed by IR scholars on the capacity of a manuscript to catalyze further inquiry, promote intellectual growth, and enhance understanding in the field. "Quality" here means that research should serve as a launching pad for additional studies—generative research that advances the state of the art, be it within the discipline at large or within more specific research programs.
- Factor 6. *Interest and topicality* underscores the importance of effective storytelling, capturing attention, and timeliness in academic research. Apart from the question of timing and an attention-grabbing style, the third item suggests that IR scholars perceive quality research as inherently linked to current, real-world events.
- Factor 7. *Challenge to existing knowledge* is a new factor that highlights the quality weight IR scholars place on questioning conventional wisdom and theoretical frameworks. Quality in those terms translates into the capacity to put forth intriguing results that defy current theories, offer evidence that disrupts existing influential ideas, and introduce general findings that prompt a critical evaluation of current knowledge. It also considers whether the work provides surprising results that nevertheless speak to existing theories, recommends changes to accepted concepts, challenges current theories, or introduces new methodologies. There is a subtle contrast in the components of factor 7, however. While the first two items and the last one highlight the importance of innovative theoretical development, the third item underscores the significance of effectively integrating novel findings into existing scholarship.

The Sociological Correlates of Quality Judgments in IR

Having outlined the cognitive dimensions of quality judgment, we now examine their associations with, and variation according to, positional variables such as gender, nationality, professional status, and epistemological stances within the discipline.

Before examining the differences based on sociological-positional variables, it is important to note that these differences may suggest that scholars identifying with these groups hold others, but also themselves, to higher quality standards than those who do not identify with these groups. Recall that while our focus is on how scholars evaluate work in IR, this does not preclude that they would hold their own work to similar standards. In fact, it is reasonable to assume that participants, to avoid cognitive dissonance, apply the same standards to themselves.

Zero-order correlations presented in table 2 show that nationality and ideology are positively associated with all seven factors. Professional rank (i.e., tenure versus untenured) shows positive associations only for tenured ranks, with moderate significance for *theoretical significance* (factor 1) and *methodological rigor* (factor 4). The correlations between gender and quality factors suggest that gender differences may be linked to variations in quality perceptions in IR, specifically in factors such as *logical style and structure* (factor 2) and *interest and topicality* (factor 6).

We regressed sociological (e.g., gender, nationality, ideology, and professional ranking/tenure status; table 3) and epistemological (e.g., theoretical affiliation and area of study; tables 4 and 5) variables on all seven quality factors to better characterize the unique associations between these positional aspects and latent quality factors.

Gender (coded for female) positively predicted *logical style and structure* (factor 2) and *interest and topicality* (factor 6)—much in line with zero-order correlations. Our findings may

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Table 2
Correlational Findings

	1	2	3	4	5	6	7	8	9	10	11	12
1. Factor 1	—											
2. Factor 2	0.30***	—										
3. Factor 3	0.54***	0.43***	—									
4. Factor 4	0.48***	0.48***	0.58***	—								
5. Factor 5	0.60***	0.42***	0.58***	0.44***	—							
6. Factor 6	0.44***	0.49***	0.59***	0.39***	0.51***	—						
7. Factor 7	0.60***	0.26***	0.53***	0.49***	0.50***	0.43***	—					
8. Gender (female)	0.01	0.10**	0.06	0.03	0.03	0.11**	−0.01	—				
9. Conservatism	0.12***	0.08*	0.25***	0.36***	0.14***	0.13***	0.20***	−0.08*	—			
10. Tenured	0.10**	0.06	0.05	0.10**	0.03	0.07*	0.08*	−0.06	0.03	—		
11. Tenure track	−0.02	−0.04	−0.03	−0.07	0.000	0.000	0.07	0.09*	0.000	−0.66***	—	
12. Nationality (South)	0.13***	0.10**	0.24***	0.20***	0.23***	0.18***	0.16***	0.05	0.30***	−0.17***	0.17***	—

Notes: Gender = male (0); female (1); other (2). Nationality = North (0); South (1).

Table 3
OLS Models for Social Variables Predicting Quality Factors

	Factor 1		Factor 2		Factor 3		Factor 4		Factor 5		Factor 6		Factor 7	
Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Tenured	0.16*	—	0.06	—	0.07	—	0.11*	—	0.08	—	0.17***	—	0.25***	—
Tenure track	0.06	—	−0.03	—	−0.03	—	−0.02	—	0.01	—	0.07	—	0.22***	—
Assistant professor	—	0.06	—	−0.03	—	−0.03	—	−0.02	—	0.01	—	0.07	—	0.22***
Associate professor	—	0.10	—	0.04	—	0.06	—	0.06	—	0.04	—	0.11*	—	0.22***
Full professor	—	0.19***	—	0.07	—	0.07	—	0.14**	—	0.11*	—	0.19***	—	0.26***
Nonacademic	—	−0.04	—	−0.01	—	0.02	—	0.01	—	0.01	—	0.04	—	0.03
Gender (female)	0.01	0.01	0.11**	0.11**	0.08*	0.08*	0.07*	0.06	0.02	0.03	0.11**	0.11**	0.00	−0.01
Nationality (South)	0.13***	0.14***	0.08*	0.09*	0.20***	0.20***	0.12***	0.14***	0.23***	0.24***	0.19***	0.17***	0.13***	0.13***
Conservatism	0.08*	0.08*	0.06	0.06	0.20***	0.20***	0.33***	0.33***	0.07	0.07	0.09*	0.09*	0.16***	0.15***
R^2	0.04	0.05	0.03	0.03	0.11	0.11	0.16	0.17	0.07	0.07	0.07	0.07	0.09	0.09
Adjusted R^2	0.04	0.04	0.02	0.02	0.11	0.10	0.15	0.16	0.06	0.06	0.06	0.06	0.08	0.08
N	789	810	789	810	789	810	789	810	789	810	789	810	789	810

Notes: Gender = male (0); female (1); other (2). Nationality = North (0); South (1). All variable values represent standardized coefficients (β). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 4
OLS Models for IR Paradigmatic Preferences Predicting Quality Factors

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Models	**(15)**	**(16)**	**(17)**	**(18)**	**(19)**	**(20)**	**(21)**
IR theory							
Realism	0.09*	0.04	0.07	0.09*	0.08*	0.05	0.04
Liberalism	0.12**	–0.01	0.07*	0.08*	0.08*	0.000	0.04
English school	0.07*	0.04	0.08*	0.09*	0.08*	0.03	0.05
Constructivism	0.13***	0.04	0.03	–0.02	0.09*	0.06	0.06
Feminism	0.02	0.01	0.01	–0.13***	0.02	0.04	–0.01
Marxism	0.08*	0.07	0.04	0.000	0.07	0.000	0.09*
Other	0.10*	0.01	0.04	–0.03	0.07	0.02	0.05
R^2	0.03	0.01	0.01	0.04	0.02	0.005	0.01
Adjusted R^2	0.02	–0.001	0.004	0.03	0.01	–0.003	0.003
N	820	820	820	820	820	820	820

Notes: All variable values represent standardized coefficients (β). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 5
OLS Models for Area of Study Predicting Quality Factors

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Models	**(22)**	**(23)**	**(24)**	**(25)**	**(26)**	**(27)**	**(28)**
Area of study							
International relations theory	0.07	0.03	–0.08	–0.12**	0.04	0.000	–0.01
International security	0.08	–0.02	0.02	–0.02	0.09*	0.02	0.000
International institutions	–0.01	0.03	0.02	–0.03	0.03	0.000	0.03
International organizations	–0.04	–0.07	–0.03	–0.08*	–0.02	–0.03	–0.05
Foreign policy analysis	0.05	0.02	0.03	0.07	0.09*	0.01	0.04
International political sociology	0.07	–0.02	0.000	–0.06	0.05	0.03	0.000
Development studies	0.08	0.01	0.08*	0.04	0.10*	0.04	0.06
International relations of a particular region	0.04	0.04	0.06	0.05	0.12**	0.12*	0.04
Other	0.02	–0.05	–0.02	–0.05	–0.03	–0.01	–0.01
R^2	0.02	0.01	0.02	0.04	0.03	0.02	0.01
Adjusted R^2	0.005	0.002	0.01	0.03	0.02	0.005	0.000
N	820	820	820	820	820	820	820

Notes: All variable values represent standardized coefficients (β). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

be taken as an expression of the quality standards female scholars hold when submitting their own work to peer review. The gender differences identified in these specific factors may reflect the gendered expectations that female scholars face in their careers, such as the additional pressure to demonstrate greater clarity and topicality in their writing compared to their male colleagues (see Leahey 2006). This interpretation of gendered differences in quality standards aligns with other studies—such as those on submission and perception gaps (Brouns 2004; Brown et al. 2020; Djupe, Smith, and Sokhe 2019; Teele and Thelen 2017)—showing that male scholars are more likely to take risks, submit more work for review, and face more rejections, while female scholars, on average, perceive themselves as less likely to be published in top general political science journals. A final note is necessary: the small effect sizes reported suggest that, despite being statistically significant, these differences may lack practical significance—particularly because previous research has found no cognitive differences between sexes (Sternberg, Wong, and Sternberg 2019; Sternberg et al. 2020).

Nationality (coded for South[13]) strongly positively predicted all seven factors. This may be an indication that scholars from the Global South hold higher quality standards than their northern peers who participated in our study. It is more likely, however, that participants from the

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Global South perceive the quality standards necessary to pass peer review to be higher than their peers in the Global North. From our data it is not possible to disentangle whether southern scholars simply view the standards as higher than their northern peers do, or whether they also harbor a sense of discrimination and unfair treatment—that they are being judged by different and higher standards than their Global North peers. This would be not surprising for factor 2, as native and non-native English speakers, by definition, face different barriers to getting published in mainstream journals as long as English is the lingua franca in academia. Scholars outside the Global North, and non-native speakers in those latitudes, are expected to have high proficiency in English (e.g., C1 level according to the Common European Framework of Reference for Languages) to pursue graduate studies, even in their own countries in some cases, and to have their academic quality recognized. As long as Anglophone norms in academic writing prevail, non-native speakers are likely to remain in a disadvantaged and structurally dependent position (Aydinli and Aydinli 2024). The discrimination and stigmatization that non-native English speakers experience in the peer review process are well documented (Demeter 2020, 31–32; Horn 2017). The increasing availability of artificial intelligence (AI) tools might help to level the playing field in terms of grammar, stylistic consistency, and even organization, flow, and readability.[14] It is worth noting, however, that because nationality shows significance across the board, the analysis of geographical differences in quality assessment goes way beyond writing and language and warrants a further study on its own.

Political orientation, measured by high scores in conservatism, emerged as a strong positive predictor of factors 3, 4, and 7. This indicates that differences in quality judgment among scholars are also channeled through ideological preferences. Our results suggest that more conservative scholars are more likely to emphasize methodological rigor and generalizability as well as practical and policy significance as markers of quality. This resonates with research in social and political psychology showing robust associations between conservatism and personality traits characterized by organization, thoroughness, productivity, and competence—that is, conscientiousness and its lower-level aspects, orderliness and industriousness (see Osborne, Satherley, and Sibley 2021).

Tenure-track rank only positively predicted *challenge to existing knowledge* (factor 7). Meanwhile, being a tenured scholar is a significant positive predictor across factors except *logical style and structure* (factor 2), *practical significance* (factor 3) and *contribution and value for future research* (factor 5). When we disaggregate professional ranking or tenure status (even-numbered models in table 3), a more nuanced picture emerges. Although all three ranks (assistant, associate, and full professorship) show significant positive associations with *challenge to existing knowledge* (factor 7), only the full professor rank predicts quality factors across the board, except for *logical style and structure* (factor 2) and *practical significance* (factor 3). Results for both groups of regression models were aligned with correlational findings. Taken together, these results suggest that early-career scholars tend to privilege, first and foremost, the challenging-existing-knowledge quality criterion. On the other hand, we observed strong and robust associations between being a full professor and valuing *theoretical significance* (factor 1) and *interest and topicality* (factor 6), which we cannot identify in the other professorial ranks. It appears that IR scholars who have progressed beyond the tenure-track stage tend to prioritize additional quality factors beyond challenging established knowledge, which is often crucial for early-career colleagues or those in the process of consolidating their positions (associate professors). Furthermore, these findings may indicate that spending more time in the profession prompts researchers to recognize and value a broader range of quality factors in scholarship.

We turn now to whether scholars' *epistemological positionality*, such as their paradigmatic affiliations and areas of study, are associated with the specific factors of quality judgment. The importance attached to the quality factors does vary according to scholars' theoretical commitment and/or subfield. Results are more nuanced than anticipated, however. The most robust positive associations between quality factors and IR theories were clustered around *theoretical significance* (factor 1) where there was a strongly significant association with constructivism, but also liberalism and other theoretical schools (except feminism). Add to this the fact that *contribution and value for future research* (factor 5) also showed an interesting contrast between mainstream and more critical theoretical approaches. Realism, liberalism, the English school, and constructivism were associated with the value-for-future-research factor, whereas feminism and Marxism did not yield any significant associations. Taken together, this can be interpreted as an indication that scholars affiliated with mainstream IR theories place a relatively higher emphasis on work that is generative within that research program, whereas scholars working within more transdisciplinary theoretical frameworks such as Marxism and feminism (and "other") view this as less important. Liberalism, realism, and the English school were also weakly positively associated with *methodological rigor* (factor 4). For liberalism and realism, this may be because this factor has a slant toward positivist methodologies (e.g., clear hypotheses, impartiality, generalizability, and empirics in the form of data), but such findings are harder to explain in the case of the English school. More noteworthy is that feminism showed a robust negative association with *methodological rigor*. This is likely due to the postpositivist epistemology and critical, normative ethos of feminist approaches to IR, which clashes with many of the items in factor 4.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

In a final exploration of affiliations to areas of study, the results are modest (table 5). The regression models yield no significant results for factors 1, 2, and 7; we are also cautious to lend much importance to the weak results for factors 3 and 6. The nuanced findings for other factors, however, unveiled some intriguing associations. Those who identified IR theory or international organization(s) as their main area of study showed a weak-to-moderate negative association with factor 4 (*methodological rigor*). Factor 5 (*contribution and value for future research*), however, coherently showed itself to be associated with subjects such as the study of the international relations of particular regions, development studies, foreign policy analysis, and international security—which in general are linked to or rooted in regional or country-based specializations.

Implications and Potential Uses

Our insights into the cognitive structures of quality judgment and their sociological aspects yield significant implications in four main areas: career and publication strategies, pedagogical and training purposes, for minorities within the discipline, and for those in positions of power.

The findings we present here can be applied to both "clinical" and "cynical" purposes (Hamati-Ataya 2012a). "Clinical" in the sense that the objectification of quality judgments generates valuable empirical knowledge about the workings of an underexamined practice in the social sciences. "Cynical" in that our mapping of the dimensions of quality judgments can be (ab)used as a strategic tool for scholars navigating the increasingly competitive landscape of contemporary academia. In an era where high-impact publishing, citation or Altmetric scores, and excellent research assessments have become more crucial than ever, understanding the factors that inform judgments of research quality is of paramount importance. Even if scholars do not agree with the factors identified, they can still be used as a reflection of how a broad cross-section of scholars assess research quality. One could use our study as a road map for meeting (or even exceeding) these often implicit standards of quality judgment more effectively and become more successful in one's publication efforts. We do not encourage such "cynical" uses of our results but believe that they are better utilized to provide transparency around quality judgments in ways that positively influence the broader discipline.

This brings us to the potential educational uses of our findings. The most straightforward application is for educators to utilize the various cognitive dimensions of quality to train students in the identified quality factors. Instructors would, of course, need to elaborate on specific quality items (e.g., what constitutes a "logical flow" or a "useful implication for theory building"), but the factors themselves offer a road map for workshops guiding students through aspects of the academic profession. For instance, some quality factors we unearthed here such as *theoretical significance* (factor 1) are arguably hard to teach, but others like the high premium placed on a clear writing style and logical organization, for instance, are more amenable to strategic maneuvering in pedagogical terms. Given the importance participants attributed to it in this study, it appears that writing has not been receiving the necessary attention in graduate training, feedback, and peer review. Another potential pedagogical application of the quality markers identified in our study involves not merely instructing students on how to utilize these markers to excel in publication outputs, but rather guiding them to reflexively engage with the process of quality judgment. The factors could, for instance, be incorporated into sessions on critical reading and peer review exercises, helping students to gain more nuanced perspectives on how they assess the quality of the work of peers and how their own work is being assessed. More importantly, our findings can prove particularly useful for raising awareness about the sociological variability of quality judgments to research students, including minority and early-career scholars, and can possibly also help to alleviate the negative effects thereof. By reflecting more on the cognitive and sociological dimension and variations in quality assessment, educational programs can play an important role in leveling these variations.

Building on this, our work may be especially beneficial for marginalized or minority scholars, who often face additional biases and barriers. The association of all quality factors with belonging to a minority or marginalized group—whether in terms of nationality, ideology, or, to some extent, gender—warrants particular attention and discussion. This suggests that scholars from these groups may either hold themselves and others to higher *individual* quality standards or perceive that the *communal* quality standards—or the barriers governing the discipline—are higher compared to those perceived by scholars belonging to the ideological, national, or gender majority. In either case, minority or marginalized groups seem to view the field as less hospitable to them and publication success as harder to achieve. This supports recent research noting that knowledge produced outside the West faces several challenges in being accepted as legitimate and valued frameworks for analyzing world politics (Chagas-Bastos 2023; 2024; Shahi 2023). Differential quality standards can constitute a real problem for a more plural, inclusive, equitable, and diverse discipline. What can be done about this and by whom? Minority scholars themselves can of course use our results to reflect on whether and how their own quality standards reflect those of majority groups and how this may or may not affect their behavior, but that burden should not of course only be on them.

Scholars in positions of power, as well as established and nonminority colleagues, may also benefit from using our work to engage in a reflexive examination of their quality

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

judgments in comparison to those of others. Here, our study of the sociological correlates of quality judgments can be put to a use that could positively influence the field. The sociological variability of quality judgments we identified can be used to inform the composition of panels aiming at maximizing diversity and representativeness in the nominations for editorial boards, conference program chairs, members of research councils, or award and prize committees. Editors or chairs of review boards could also use these findings to "review the reviewers," providing the peer review process with a transparent baseline for assessing whether a given review emphasizes conventional, more idiosyncratic, or sociointellectually specific quality factors in a manuscript. Directive boards in research institutions and funding bodies could also leverage our findings to develop policies and guidelines based on specific cognitive aspects they deem important in the material they should judge, aimed at ensuring fairer evaluations across the board and mitigating personal "tastes" and cognitive variations in how reviewers and colleagues in positions of power understand such guidelines. Such transparency about quality judgment could be generally beneficial for the field, but perhaps especially for early-career and marginalized scholars.

Conclusion

This article presents a novel and systematic investigation into the cognitive dimensions of quality judgments and how these factors vary according to sociological correlates. We go beyond output- and process-oriented approaches to studying research quality by shifting the focus from why some papers accrue more citations and the biases in this process to a deeper and systematic understanding of how scholars evaluate research quality. Incorporating behavioral science into the sociology of knowledge opens a promising interdisciplinary research avenue for examining how knowledge is produced, validated, and shared in different social contexts. It brings new evidence to previous research on quality judgments that highlighted the role of emotional and moral factors in scholarly judgment of "excellent" research (Lamont 2009, chap. 6). Despite these psychological aspects, the influence of personal tastes, disciplinary ethos, and praxis on evaluations of research quality, which may vary according to one's positionality—whether in terms of gender, nationality, or political orientation—we demonstrate that quality judgment is grounded in deep-seated, common, and shared cognitive stances.

Our study also lays the groundwork for considering broader implications and future research avenues. It is noteworthy that we replicated the six-factor structure identified in the original study by Sternberg and Gordeeva (1996), even when applying the measurement to a different disciplinary field and within a different temporal context. The consistently high Cronbach's α coefficients —demonstrating strong reliability in our scale, despite minor changes to some items and the addition of five new ones—suggest a level of resilience that warrants further investigation. Future research should, for example, test the psychometric properties of our scale and seek to replicate our findings across different disciplines adopting the same cross-national approach employed in our study.

That said, a plethora of quality items other than the ones explored here could be conceived. Interdisciplinarity, which is increasingly vital in addressing societal challenges, for instance, does not fit neatly into the items surveyed and may even stand in opposition to some of the more discipline-oriented items surveyed. Reflexivity could also be an obvious candidate. This would be a possible quality item transcending the equation of methodological quality with impartiality, technical proficiency, and rigor in data collection and analysis, embracing a more holistic view where the researcher's positionality and the research process itself are scrutinized for their integrity and ethical soundness. In fact, the political and ethical positioning of a research output—whether the research leading to it has been conducted and applied ethically, without exploiting, misrepresenting, or causing harm—may be another related quality criterion that escaped our study. It should also consider its broader impact on society and specific communities, prioritizing the welfare and rights of participants and disadvantaged groups. The list continues, and we encourage further researchers to build upon the study of quality judgment.

It would also be interesting to expand the research on the products under judgment. Given the myriad of research outputs, we have reason to believe that researchers may apply the set of factors we unearthed in our study differently depending on the type of scholarly work at hand, whether it is a research paper, a monograph, or a research proposal submitted to funding agencies. These contexts of quality judgment other than peer-reviewed research articles may also intersect in different ways with positionality variables, and further studies could inquire into these context dependencies. Furthermore, a qualitative investigation of our findings on how marginalized scholars tend to view quality standards as consistently higher would also be highly relevant.

Lastly, our interrogation of geographical positionality only scratched the surface. Geoepistemic locations contain enormous differences, contradictions, and complexities, and it would be interesting to explore these with more fine-grained analyses. Nationality alone should not account for the ways in which geographical positionality may affect quality judgment. Further research should explore whether nationality, institutional affiliation, or educational environment most significantly impact quality judgments, as well as the interactions between these variables, such as career mobility and socialization. These elements may influence how quality is perceived and judged, given the scholarly incentives in a particular national context, institutional and

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

working conditions, access to resources and funding, and other related aspects.

All in all, we offer an initial exploration of these questions, but we believe they could pave the way for research aimed at developing a more robust and transparent understanding of the psychological and social factors that underpin the criteria used by scholarly communities to define quality research. In the long term, we hope our approach leads to a broader understanding of academic excellence and promotes more inclusive standards of quality in research.

Supplementary material

To view supplementary material for this article, please visit http://doi.org/10.1017/S1537592724002676.

Data replication

Data replication sets are available in Harvard Dataverse at: https://doi.org/10.7910/DVN/E2DSBP.

Acknowledgments

This project has received generous funding from Horizon Europe's Marie Skłodowska-Curie Actions (grant agreement no. 101032425). We extend our gratitude to Michelle Dion, Kristin Eggeling, Yong-Soo Eun, Robert Sternberg, Benjamin de Carvalho, Halvard Leira, Kyle Grayson, William Wohlforth, the editors, and anonymous reviewers, as well as the members of the IR Group at the University of Copenhagen for their invaluable comments and suggestions. Special thanks are due to Lisa Verzier for her dedicated research support.

Notes

1 Potential exceptions are the early iterations of the Teaching, Research and International Policy (TRIP) survey, which asked participants to list the most influential scholars and journals in the IR field. While these surveys offer some insight into how segments of the IR scholarly community rank *influence*, they neither examine quality nor the underlying processes of such judgments, as we do here.

2 Sternberg (2011) recently expanded his original model by incorporating *wisdom*, which he defines as the ethical use of one's abilities and knowledge to achieve the common good.

3 Cognitive style denotes enduring attitudes, preferences, or customary strategies that shape an individual's way of perceiving, remembering, thinking, learning, and problem solving. It facilitates adaptation to the external world, evolving through interaction with the surrounding environment (Kozhevnikov 2007; Sternberg 1997).

4 The 2017 wave of the TRIP survey provides some insight into the potential size of the global IR community, reporting a sample of 13,482 scholars from 36 countries (Maliniak et al. 2017).

5 `R` package `pwr` (Champely et al. 2020).

6 Coefficient conversion was set according to Cohen's $f^2 = 0.10$ for small effects, which equates to regression standardized coefficients ranging from $\beta = 0.04$ to $\beta = 0.25$ (Cohen 1988).

7 Whether IR is considered a subfield of political science or a discipline in its own right varies depending on institutional differences around the world.

8 We ended with a list of 77,897 email addresses. After an automated process to remove duplicate addresses—carefully avoiding the exclusion of homonyms—our database was reduced to 77,701 unique entries. To improve success rates in attracting participants, we employed an email sequence, with follow-up emails sent every fortnight to reengage potential participants. After completing all three sequences, 15.74% of the emails (soft or hard) bounced, equating to approximately 12,227 undelivered emails. In the end, we successfully contacted 65,474 recipients, with 46.8% (30,642 recipients) opening our invitation to participate in the study.

9 The questions referring to theoretical preferences and areas of study were taken from the 2017 TRIP Faculty Survey (Maliniak et al. 2017).

10 Basic statistics on item ratings for the quality questionnaire (research quality scale) can be found in appendix D of the online supplementary file.

11 It is important to note that factors are not ranked by the importance respondents ascribe to their items —i.e., the mean scale value—but by their ability to account for the variance in the observed data.

12 Developing new methodologies is also included in this factor, although the respective item is the weakest loading onto the factor.

13 The variable *nationality* is a dummy variable (0 = North; 1 = South) aggregating self-reported country-of-nationality data. The North–South coding is based on the United Nations (2019) list of developed economies, combined with the International Monetary Fund's (2023) list of developed economies. Countries that are considered part of the "South" or "non-West" in the global IR literature due to their distance from the discipline's mainstream, such as Japan, Hong Kong, Israel, South Korea, Singapore, and Taiwan, were then categorized as northern countries. We acknowledge, however, that "North" and "South" are no longer solely economic categories, but also may be used to express political and historical subjectivities.

14 AI-based tools may also, however, increase inequalities in the academic profession, depending on how editorial policies for the use of AI are established.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

References

Aksnes, Dag W. 2003. "Characteristics of Highly Cited Papers." *Research Evaluation* 12 (3): 159–70. DOI: 10.3152/147154403781776645.

Alter, Karen J., Jean Clipperton, Emily Schraudenbach, and Laura Rozier. 2020. "Gender and Status in American Political Science: Who Determines Whether a Scholar Is Noteworthy?" *Perspectives on Politics* 18 (4): 1048–67. DOI: 10.1017/s1537592719004985.

Avey, Paul C., and Michael C. Desch. 2014. "What Do Policymakers Want from Us? Results of a Survey of Current and Former Senior National Security Decision Makers." *International Studies Quarterly* 58 (2): 227–46. DOI: 10.1111/isqu.12111.

Avey, Paul C., Michael C. Desch, Eric Parajon, Susan Peterson, Ryan Powers, and Michael J. Tierney. 2022. "Does Social Science Inform Foreign Policy? Evidence from a Survey of US National Security, Trade, and Development Officials." *International Studies Quarterly* 66 (1): sqab057. DOI: 10.1093/isq/sqab057.

Aydinli, Ersel, and Julie Mathews. 2000. "Are the Core and Periphery Irreconcilable? The Curious World of Publishing in Contemporary International Relations." *International Studies Perspectives* 1 (3): 289–303. DOI: 10.1111/1528-3577.00028.

Aydinli, Ersel, and Julie Aydinli. 2024. "Exposing Linguistic Imperialism: Why Global IR Has to Be Multilingual." *Review of International Studies* 50 (6): 943–64. DOI: 10.1017/s0260210523000700.

Baele, Stephane, and Gregorio Bettiza. 2021. "'Turning' Everywhere in IR: on the Sociological Underpinnings of the Field's Proliferating Turns." *International Theory* 13 (2): 314–40. DOI: 10.1017/s1752971920000172.

Baldi, Stephane. 1998. "Normative versus Social Constructivist Processes in the Allocation of Citations: A Network-Analytic Model." *American Sociological Review* 63 (6): 829–46. DOI: 10.2307/2657504.

Breitenbauch, Henrik. 2013. *International Relations in France: Writing between Discipline and State*. New York: Routledge. DOI: 10.4324/9780203403167.

Brouns, Margo. 2004. "Gender and the Assessment of Scientific Quality." In *Gender and Excellence in the Making*, eds. Delilah Al-Khudhairy, Nicole Dewandre, and Helen Wallace, 147–54. Community Research Study EUR 21222. Luxembourg: Office for Official Publications of the European Communities.

Brown, Nadia E., Yusaku Horiuchi, Mala Htun, and David Samuels. 2020 "Gender Gaps in Perceptions of Political Science Journals." *PS: Political Science & Politics* 53 (1): 114–21. DOI: 10.1017/s1049096519001227.

Bryant, Fred, and Paul Yarnold. 1995. "Principal Components Analysis and Exploratory and Confirmatory Factor Analysis." In *Reading and Understanding Multivariate Statistics*, eds. L. G. Grimm and R. R. Yarnold, 99–136. Washington, DC: American Psychological Association.

Chagas-Bastos, Fabrício H. 2023. "International Insertion: A Non-Western Contribution to International Relations." In *Oxford Research Encyclopedia of International Studies*, eds. Nukhet Sandal, Victor Asal, Gilbert M. Khadiagala, Nalanda Roy, Cintia Quiliconi, and Matthew Weinert. Oxford: Oxford University Press. DOI: 10.1093/acrefore/9780190846626.013.652.

———. 2024. "The Challenge for the 'Rest': Insertion, Agency Spaces and Recognition in World Politics." *International Affairs* 100 (1): 43–60. DOI: 10.1093/ia/iiad246.

Chagas-Bastos, Fabrício H., Erica S. Resende, Faten Ghosn, and Debbie Lisle. 2023. "Navigating the Global South Landscape: Insights and Implications for Representation and Inclusion in ISA Journals." *International Studies Perspectives* 24 (4): 441–66. DOI: 10.1093/isp/ekad010.

Chagas-Bastos, Fabrício H., and Peter Marcus Kristensen. 2025. "Replication Data for: Mapping Quality Judgment in International Relations: Cognitive Dimensions and Sociological Correlates." *Harvard Dataverse*. DOI: 10.7910/DVN/E2DSBP.

Champely, Stephan, Claus Ekstrom, Peter Dalgaard, Jeffrey Gill, Stephan Weibelzahl, Aditya Anandkumar, Clay Ford, Robert Volcic, and Helios De Rosario. 2020. `pwr`: Basic Functions for Power Analysis. Package version 1.3-0, March 17. Vienna: Cran.R Project. https://cran.r-project.org/web/packages/pwr.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum.

Colgan, Jeff D. 2016. "Where Is International Relations Going? Evidence from Graduate Training." *International Studies Quarterly* 60 (3): 486–98. DOI: 10.1093/isq/sqv017.

Comrey, Andrew L., and Howard B. Lee. 1992. *A First Course in Factor Analysis*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum.

Darwich, May, Morten Valbjørn, Bassel F. Salloukh, Waleed Hazbun, Amira Abu Samra, Said Saddiki, Adham Saouli, Hamad H. Albloshi, and Karim Makdisi. 2021. "The Politics of Teaching International Relations in the Arab World: Reading Walt in Beirut, Wendt in Doha, and Abul-Fadl in Cairo." *International Studies Perspectives* 22 (4): 407–38. DOI: 10.1093/isp/ekaa020.

Demeter, Márton. 2020. *Academic Knowledge Production and the Global South: Questioning Inequality and Under-Representation*. Cham: Palgrave Macmillan. DOI: 10.1007/978-3-030-52701-3.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Desch, Michael. 2019. *Cult of the Irrelevant: The Waning Influence of Social Science on National Security*. Princeton, NJ: Princeton University Press. DOI: 10.23943/princeton/9780691181219.001.0001.
Dion, Michelle L., Jane Lawrence Sumner, and Sara McLaughlin Mitchell. 2018. "Gendered Citation Patterns across Political Science and Social Science Methodology Fields." *Political Analysis* 26 (3): 312–27. DOI: 10.1017/pan.2018.12.
Djupe, Paul A., Amy E. Smith, and Anand Edward Sokhe. 2019. "Explaining Gender in the Journals: How Submission Practices Affect Publication Patterns in Political Science." *PS: Political Science & Politics* 52 (1): 71–77. DOI: 10.1017/s104909651800104x.
Dunne, Tim, Lene Hansen, and Colin Wight. 2013. "The End of International Relations Theory?" *European Journal of International Relations* 19 (3): 405–25. DOI: 10.1177/1354066113495485.
Entringer, Irene, Shauna N. Gillooly, Susan Peterson, Ryan Powers, and Michael J. Tierney. 2023. "TRIP Faculty Survey 2022–2023 Report." Teaching, Research, and International Policy Project, March. Williamsburg, VA: College of William & Mary. https://trip.wm.edu/research/faculty-surveys/Faculty-Survey-2022-Partial-Report.pdf.
Ettinger, Aaron. 2020. "Scattered and Unsystematic: The Taught Discipline in the Intellectual Life of International Relations." *International Studies Perspectives* 21 (3): 338–61. DOI: 10.1093/isp/ekz028.
———. 2023. "Global International Relations and Worlding beyond the West: A Pedagogical Critique." *International Studies Review* 25 (4): viad052. DOI: 10.1093/isr/viad052.
Everett, Jim A. C. 2013. "The 12 Item Social and Economic Conservatism Scale (SECS)." *PloS ONE* 8 (12): e82131. DOI: 10.1371/journal.pone.0082131.
Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. 1994. *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: SAGE.
Gofas, Andreas, Inanna Hamati-Ataya, and Nicholas Onuf, eds. 2018. *The SAGE Handbook of the History, Philosophy and Sociology of International Relations*. London: SAGE. DOI: 10.4135/9781526402066.
Goh, Evelyn. 2019. "US Dominance and American Bias in International Relations Scholarship: A View from the Outside." *Journal of Global Security Studies* 4 (3): 402–10. DOI: 10.1093/jogss/ogz029.
Gorsuch, Richard L. 2015. *Factor Analysis*, 2nd edition. New York: Routledge. DOI: 10.4324/9780203781098.
Gottfredson, Stephen. 1978. "Evaluating Psychological Research Reports: Dimensions, Reliability, and Correlates of Quality Judgments." *American Psychologist* 33 (10): 920–34. DOI: 10.1037/0003-066X.33.10.920.
Grenier, Félix, and Jonas Hagmann. 2016. "Sites of Knowledge (Re-)Production: Toward an Institutional Sociology of International Relations Scholarship." *International Studies Review* 18 (2): 333–65. DOI: 10.1093/isr/viw006.
Grenier, Félix, Jonas Hagmann, Thomas Biersteker, Marina Lebedeva, Yulia Nikitina, and Ekaterina Koldunova. 2020. "The Institutional 'Hinge': How the End of the Cold War Conditioned Canadian, Russian, and Swiss IR Scholarship." *International Studies Perspectives* 21 (2): 198–217. DOI: 10.1093/isp/ekz021.
Guetzkow, Joshua, Michèle Lamont, and Grégoire Mallard. 2004. "What Is Originality in the Humanities and the Social Sciences?" *American Sociological Review* 69 (2): 190–212. DOI: 10.1177/000312240406900203.
Guzzini, Stefano. 2013. "The Ends of International Relations Theory: Stages of Reflexivity and Modes of Theorizing." *European Journal of International Relations* 19 (3): 521–41. DOI: 10.1177/1354066113494327.
Hagmann, Jonas, and Thomas J. Biersteker. 2014. "Beyond the Published Discipline: Toward a Critical Pedagogy of International Studies." *European Journal of International Relations* 20 (2): 291–315. DOI: 10.1177/1354066112449879.
Hamati-Ataya, Inanna. 2012a. "IR Theory as International Practice/Agency: A Clinical-Cynical Bourdieusian Perspective." *Millennium: Journal of International Studies* 40 (3): 625–46. DOI: 10.1177/0305829812442234.
———. 2012b. "Reflectivity, Reflexivity, Reflexivism: IR's 'Reflexive Turn'—and Beyond." *European Journal of International Relations* 19 (4): 669–94. DOI: 10.1177/1354066112437770.
Haslam, Nick, Lauren Ban, Leah Kaufmann, Stephen Loughnan, Kim Peters, Jennifer Whelan, and Sam Wilson. 2008. "What Makes an Article Influential? Predicting Impact in Social and Personality Psychology." *Scientometrics* 76 (1): 169–85. DOI: 10.1007/s11192-007-1892-8.
Heiskanen, Jaakko, and Paul Beaumont. 2024. "Reflex to Turn: The Rise of Turn-Talk in International Relations." *European Journal of International Relations* 30 (1): 3–26. DOI: 10.1177/13540661231205694.
Hendrix, Cullen S., and Jon Vreede. 2019. "US Dominance in International Relations and Security Scholarship in Leading Journals." *Journal of Global Security Studies* 4 (3): 310–20. DOI: 10.1093/jogss/ogz023.
Hendrix, Cullen S., Julia Macdonald, Ryan Powers, Susan Peterson, and Michael J. Tierney. 2023. "The Cult of the Relevant: International Relations Scholars and Policy Engagement beyond the Ivory Tower."

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Perspectives on Politics 21 (4): 1270–82. DOI: 10.1017/s153759272300035x.
Herrmannova, Drahomira, Robert M. Patton, Petr Knoth, and Christopher G. Stahl. 2018. "Do Citations and Readership Identify Seminal Publications?" *Scientometrics* 115 (1): 239–62. DOI: 10.1007/s11192-018-2669-y.
Horn, Sierk. 2017. "Non-English Nativeness as Stigma in Academic Settings." *Academy of Management Learning & Education* 16 (4): 579–602. DOI: 10.5465/amle.2015.0194.
Hvid, Aksel, Fabrício H. Chagas-Bastos, and Peter Marcus Kristensen. 2025. "Power Shifts and Knowledge Production: India's Rise and Scholarship in International Relations." *All Azimuth* (forthcoming).
International Monetary Fund. 2023. "World Economic Outlook Database: Groups and Aggregates Information." Country Composition of WEO Groups, updated April 2023. https://www.imf.org/en/Publications/WEO/weo-database/2023/April/groups-and-aggregates. Accessed March 16, 2024.
Kozhevnikov, Maria. 2007. "Cognitive Styles in the Context of Modern Psychology: Toward an Integrated Framework of Cognitive Style." *Psychological Bulletin* 133 (3): 464–81. DOI: 10.1037/0033-2909.133.3.464.
Kristensen, Peter Marcus. 2012. "Dividing Discipline: Structures of Communication in International Relations." *International Studies Review* 14 (1): 32–50. DOI: 10.1111/j.1468-2486.2012.01101.x.
———. 2015. "Revisiting the 'American Social Science'—Mapping the Geography of International Relations." *International Studies Perspectives* 16 (3): 246–69. DOI: 10.1111/insp.12061.
———. 2018. "International Relations at the End: A Sociological Autopsy." *International Studies Quarterly* 62 (2): 245–59. DOI: 10.1093/isq/sqy002.
Kristensen, Peter Marcus, and Ras Tind Nielsen. 2010. "Writing on the Wall: Prominence, Promotion, Power Politics and the Innovation of a Chinese International Relations Theory." Master's thesis. University of Copenhagen.
Lamont, Michèle. 1987. "How to Become a Dominant French Philosopher: The Case of Jacques Derrida." *American Journal of Sociology* 93 (3): 584–622. DOI: 10.1086/228790.
———. 2009. *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press. DOI: 10.4159/9780674054158.
Leahey, Erin. 2006. "Gender Differences in Productivity: Research Specialization as a Missing Link." *Gender and Society* 20 (6): 754–80. DOI: 10.1177/0891243206293030.
Lee, John D., Kim J. Vicente, Andrea Cassano, and Anna Shearer. 2003. "Can Scientific Impact Be Judged Prospectively? A Bibliometric Test of Simonton's Model of Creative Productivity." *Scientometrics* 56 (2): 223–32. DOI: 10.1023/A:1021967111530.
Lohaus, Mathis, and Wiebke Wemheuer-Vogelaar. 2021. "Who Publishes Where? Exploring the Geographic Diversity of Global IR Journals." *International Studies Review* 23 (3): 645–69. DOI: 10.1093/isr/viaa062.
Maliniak, Daniel, Ryan Powers, and Barbara F. Walter. 2013. "The Gender Citation Gap in International Relations." *International Organization* 67 (4): 889–922. DOI: 10.1017/s0020818313000209.
Maliniak, Daniel, Susan Peterson, and Michael J. Tierney. 2012. "TRIP around the World: Teaching, Research, and Policy Views of International Relations Faculty in 20 Countries." Teaching, Research, and International Policy Project. Williamsburg, VA: College of William & Mary. https://www.wm.edu/offices/global-research/_documents/trip/trip_around_the_world_2011.pdf.
Maliniak, Daniel, Susan Peterson, Ryan Powers, and Michael J. Tierney. 2017. "TRIP 2017 Faculty Survey." Teaching, Research, and International Policy Project. Williamsburg, VA: College of William & Mary. https://trip.wm.edu/research/faculty-surveys.
———. 2018. "Is International Relations a Global Discipline? Hegemony, Insularity, and Diversity in the Field." *Security Studies* 27 (3): 448–84. DOI: 10.1080/09636412.2017.1416824.
Mallard, Grégoire, Michèle Lamont, and Joshua Guetzkow. 2009. "Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review." *Science, Technology, & Human Values* 34 (5): 573–606. DOI: 10.1177/0162243908329381.
Mearsheimer, John J., and Stephen M. Walt. 2013. "Leaving Theory Behind: Why Simplistic Hypothesis Testing Is Bad for International Relations." *European Journal of International Relations* 19 (3): 427–57. DOI: 10.1177/1354066113494320.
Mitchell, Sara McLaughlin, Samantha Lange, and Holly Brus. 2013. "Gendered Citation Patterns in International Relations Journals." *International Studies Perspectives* 14 (4): 485–92. DOI: 10.1111/insp.12026.
Murphy, Michael, Andrew Heffernan, Caroline Dunton, and Amelia C. Arsenault. 2023. "The Disciplinary Scholarship of Teaching and Learning in Political Science and International Relations: Methods, Topics, and Impact." *International Politics* 60 (5): 1030–48. DOI: 10.1057/s41311-022-00425-5.
Nowotny, Helga, Peter Scott, and Michael Gibbons. 2003. "Introduction: 'Mode 2' Revisited: The New Production of Knowledge." *Minerva* 41 (3): 179–94. https://www.jstor.org/stable/41821245.
Osborne, Danny, Nicole Satherley, and Chris G. Sibley. 2021. "Personality and Ideology: A Meta-Analysis of the Reliable, but Non-Causal, Association between

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Openness and Conservatism." In *The Oxford Handbook of Behavioral Political Science*, eds. Alex Mintz and Lesley G. Terris, 315–56. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780190634131.013.35.
Østby, Gudrun, Håvard Strand, Ragnhild Nordås, and Nils Petter Gleditsch. 2013. "Gender Gap or Gender Bias in Peace Research? Publication Patterns and Citation Rates for *Journal of Peace Research*, 1983–2008." *International Studies Perspectives* 14 (4): 493–506. DOI: 10.1111/insp.12025.
Phull, Kiran, Gokhan Ciflikli, Gustav Meibauer. 2019. "Gender and Bias in the International Relations Curriculum: Insights from Reading Lists." *European Journal of International Relations* 25 (2): 383–407. DOI: 10.1177/1354066118791690.
Rathbun, Brian. 2012. "Politics and Paradigm Preferences: The Implicit Ideology of International Relations Scholars." *International Studies Quarterly* 56 (3): 607–22. DOI: 10.1111/j.1468-2478.2012.00749.x.
Russett, Bruce, and Taylor Arnold. 2010. "Who Talks, and Who's Listening? Networks of International Security Studies." *Security Dialogue* 41 (6): 589–98. DOI: 10.1177/0967010610388205.
Seabrooke, Leonard, and Kevin L. Young. 2017. "The Networks and Niches of International Political Economy." *Review of International Political Economy* 24 (2): 288–331. DOI: 10.1080/09692290.2016.1276949.
Shadish, William R., Jr. 1989. "The Perception and Evaluation of Quality in Science." In *The Psychology of Science: Contributions to Metascience*, eds. Barry Gholson, William R. Shadish Jr., Robert A. Neimeyer, and Arthur C. Houts, 383–436. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9781139173667.021.
Shahi, Deepshikha. 2023. *Global IR Research Programme: The Futuristic Foundation of "One and Many."* Cham: Palgrave Macmillan. DOI: 10.1007/978-3-031-39121-7.
Sillanpää, Antti, and Tommi Koivula. 2010. "Mapping Conflict Research: A Bibliometric Study of Contemporary Scientific Discourses." *International Studies Perspectives* 11 (2): 148–71. DOI: 10.1111/j.1528-3585.2010.00399.x.
Soreanu, Raluca, and David Hudson. 2008. "Feminist Scholarship in International Relations and the Politics of Disciplinary Emotion." *Millennium: Journal of International Studies* 37 (1): 123–51. DOI: 10.1177/0305829808093768.
Sternberg, Robert J. 1985a. *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge: Cambridge University Press.
——. 1985b. "Implicit Theories of Intelligence, Creativity, and Wisdom." *Journal of Personality and Social Psychology* 49 (3): 607–27. DOI: 10.1037//0022-3514.49.3.607.
—— 1997. *Thinking Styles*. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511584152.
—— 2011. "The Theory of Successful Intelligence." In *Cambridge Handbook of Intelligence*, eds. Robert J. Sternberg and Scott Barry Kaufman, 504–27. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511977244.026.
Sternberg, Robert J., Chak Haang Wong, and Karin Sternberg. 2019. "The Relation of Tests of Scientific Reasoning to Each Other and to Tests of General Intelligence." *Journal of Intelligence* 7 (3). DOI: 10.3390/jintelligence7030020.
Sternberg, Robert, Elena L. Grigorenko, Michel Ferrari, and Pamela Clinkenbeard. 1999. "A Triarchic Analysis of an Aptitude-Treatment Interaction." *European Journal of Psychological Assessment* 15 (1): 3–13. DOI: 10.1027//1015-5759.15.1.3.
Sternberg, Robert, Rebel J. E. Todhunter, Aaron Litvak, and Karin Sternberg. 2020. "The Relation of Scientific Creativity and Evaluation of Scientific Impact to Scientific Reasoning and General Intelligence." *Journal of Intelligence* 8 (2): 17. DOI: 10.3390/jintelligence8020017.
Sternberg, Robert J., and Tamara Gordeeva. 1996. "The Anatomy of Impact: What Makes an Article Influential?" *Psychological Science* 7 (2): 69–75. DOI: 10.1111/j.1467-9280.1996.tb00332.x.
Strang, David, and Kyle Siler. 2015. "Revising as Reframing: Original Submissions versus Published Papers in *Administrative Science Quarterly*, 2005 to 2009." *Sociological Theory* 33 (1): 71–96. DOI: 10.1177/0735275115572152.
Sylvester, Christine. 2007. "Whither the International at the End of IR." *Millennium: Journal of International Studies* 35 (3): 551–73. DOI: 10.1177/03058298070350031101.
Teele, Dawn Langan, and Kathleen Thelen. 2017. "Gender in the Journals: Publication Patterns in Political Science." *PS: Political Science & Politics* 50 (2): 433–47. DOI: 10.1017/s1049096516002985.
United Nations. 2019. *World Economic Situation and Prospects Report 2019*. New York: United Nations. https://www.un-ilibrary.org/content/books/9789210476119. Accessed March 16, 2024.
Wæver, Ole. 1998. "The Sociology of a Not So International Discipline: American and European Developments in International Relations." *International Organization* 52 (4): 687–727. DOI: 10.1162/002081898550725.
——. 2016. "Still a Discipline after All These Debates?" In *International Relations Theories*, 4th edition, eds. Tim Dunne, Milja Kurki, and Steve Smith, 322–45. Oxford: Oxford University Press. DOI: 10.1093/hepl/9780198707561.003.0017.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press

Walt, Stephen M. 1999. "Rigor or Rigor Mortis? Rational Choice and Security Studies." *International Security* 23 (4): 5–48. DOI: 10.1162/isec.23.4.5.

Wemheuer-Vogelaar, Wiebke, Peter Marcus Kristensen, and Mathis Lohaus. 2022. "The Global Division of Labor in a Not So Global Discipline." *All Azimuth* 11 (1): 3–27. DOI: 10.20991/allazimuth.1034358.

Xia, Wanjun, Tianrui Li, and Chongshou Li. 2023. "A Review of Scientific Impact Prediction: Tasks, Features and Methods." *Scientometrics* 128: 543–85. DOI: 10.1007/s11192-022-04547-8.

https://doi.org/10.1017/S1537592724002676 Published online by Cambridge University Press