RESEARCH ARTICLE

Against method: Exploding the boundary between qualitative and quantitative studies of science

Donghyun Kang[1] and James Evans[1,2]

[1]Sociology Department, University of Chicago
[2]Santa Fe Institute

an open access journal

Check for updates

Citation: Kang, D., & Evans, J. (2020). Against method: Exploding the boundary between qualitative and quantitative studies of science. *Quantitative Science Studies*, *1*(3), 930–944. https://doi.org/10.1162/qss_a_00056

DOI:
https://doi.org/10.1162/qss_a_00056

Corresponding Author:
James Evans
jevans@uchicago.edu

Handling Editors:
Loet Leydesdorff, Ismael Rafols, and Staša Milojević

Keywords: content analysis, mixed methods, qualitative analysis, quantitative analysis, science studies, word embedding

ABSTRACT

Quantitative and qualitative studies of science have historically played radically different roles with opposing epistemological commitments. Using large-scale text analysis, we see that qualitative studies generate and value new theory, especially regarding the complex social and political contexts of scientific action, while quantitative approaches confirm existing theory and evaluate the performance of scientific institutions. Large-scale digital data and emerging computational methods could allow us to refigure these positions, turning qualitative artifacts into quantitative patterns into qualitative insights across many scales, heralding a new era of theory development, engagement, and relevance for scientists, policy-makers, and society.

1. INTRODUCTION

In Paul Feyerabend's philosophical treatise *Against Method* (1975), he outlined an "anarchistic theory of knowledge" that argued all major scientific advances eschew any generalizable notion of scientific method. Methodological innovations and not tradition portend punctuated progress. He held that prohibitions against *ad hoc* hypotheses and inconsistent findings, along with a focus on theoretical falsification, decreased the potential for new discoveries and insights. These efforts to formalize and systematize science also drastically limit the potential for fields to learn from one another. In this article, we empirically examine the vast and growing divide between quantitative and qualitative studies of science.

Here we show that quantitative and qualitative science studies represent not only distinctive objects and approaches of study—distinctive ontologies and epistemologies—but that they manifest diametrically opposed evaluations of the same objects and approaches. Recent research on scientific review reveals that researchers further apart in networks of collaboration are more likely to dispute the validity of those investigations (Teplitskiy, Acuna, et al., 2018). The divide between quantitative and qualitative studies of science is deeper, and we will argue this divide limits the potential for insight, advance and relevance that could come from greater collaboration and an explosion of ontological and epistemic commitments.

British scientist and novelist C. P. Snow argued, in an influential 1959 Rede Lecture, that Western intellectual life was divided into "two cultures" (Snow, 1959)—sciences and humanities. He memorably ridiculed the British educational establishment for overrewarding the humanities at the expense of scientific literacy. "Once or twice I have been provoked [by snobbish literary society] and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold: it was also negative. Yet I was asking … the scientific equivalent of: Have you read a work of Shakespeare's?" The thrust of Snow's argument is not that science is better than humanistic insight, but that intelligent inquiry is divided and that this division may represent the most prominent barrier to recognizing, grappling with, and ultimately solving the world's problems[1]. We will perform a 21st-century replication of Snow's cocktail party experiment on quantitative and qualitative students of science, and argue forcefully that a reformulated treaty between quantitative and qualitative studies could portend a renaissance of engagement with scientists about what they do—forward and defend supportable claims—at the level at which they do it. It could also span scales of analysis, connecting scientific action to science policy and civic action, leading to new discoveries and new relevance. For the prospect of more deeply understanding and engaging with science, we argue against method.

Prior research has demonstrated important differences between quantitative and qualitative studies of science. A recent *Journal of Informetrics* analysis of qualitative versus quantitative science studies handbooks produced between 1977 and 2008 suggested that these areas tend to discuss different ideas, characterized by distinctive words, which cluster in very different areas of a high-dimensional "word space" (Milojević, Sugimoto, et al., 2014). Moreover, when the same authors analyzed articles from a sample of quantitative and qualitative journals similar to those we explore below, they found a related clustering pattern. Leydesdorff and Van Den Besselaar (1997) rendered articles from the same sets of journals into a comparable "citation space," which revealed comparable clustering and demonstrated the limited degree to which quantitative and qualitative science studies drew from and referenced one another.

Our goal here is to revisit the comparison of quantitative and qualitative science studies with emerging tools from machine learning that allow us to survey not only the distinctive semantic focus and approach in these areas but also their distinctive evaluation of the same. As we reveal below, this demonstrates not only difference but direct opposition between quantitative and qualitative ontologies and epistemologies that suggest the potential for radical complementarity. The strengths of one are the weaknesses of the other, posing powerful opportunities for synthesis.

We begin by detailing our empirical investigation of qualitative and quantitative studies of science, articulating and interpreting distinctions between these approaches. Then we detail how new data on science and computational methods open up new pathways for collaboration and mutual learning that could provoke advance in our understanding and engagement with science.

2. INTERROGATING QUALITATIVE AND QUANTITATIVE SCIENCE STUDIES

We collected titles and abstracts from five journals. From the qualitative side, we selected *Social Studies of Science* (*SSS*), *Science, Technology, & Human Values* (*STHV*), and *Minerva*; from the quantitative side, we considered *Scientometrics* (*SCI*) and *Research Policy* (*RP*). We note that Milojević et al. (2014) initially considered articles from *RP* indeterminately quantitative or qualitative, but their analysis demonstrated that more than 80% of those articles could be classified as quantitative studies. The same logic could be applied for *SSS*, as approximately 10% of its articles were quantitative, according to their classification. We anchored our categories with classifications of the journals themselves, as even quantitative studies in *SSS* and qualitative studies in *RP* reference other work from within the same journal and typically reflect epistemological commitments there.

We first collected sets of digital object identifiers (DOI) from each journal. We used the Springer API for *SCI* and *Minerva*, and the Elsevier API for *RP*, to retrieve all DOIs affiliated with the journals. Subsequently, titles and abstracts were also retrieved using the APIs. For *SSS* and *STHV*, DOIs were

[1] Snow argued that because scientists felt humanistic inquiry irrelevant, "their imaginative understanding is less than it could be. They are self-impoverished."

Table 1. Data involved in analysis

Category	Journal	Journal start	Collection start	DOIs	Documents with abstracts	DOIs/Abstracts
Qualitative	*Science, Technology and Human Values (STHV)*	1988[a]	1988	1,180	838	71%
	Social Studies of Science (SSS)	1971	1978	1,916	1,151	60%
	Minerva	1962	1965	1,899	396	21%[b]
Subtotal				**4,995**	**2,385**	**48%**
Quantitative	*Scientometrics (SCI)*	1978	1978	5,953	5,383	90%
	Research Policy (RP)	1971	1971	4,031	3,234	80%
Subtotal				**9,984**	**8,617**	**86%**
Total				**14,979**	**11,002**	**73%**

[a] We decided to only take into account articles published *STHV* from 1988, considering its institutional shift to 4S, when the journal started to rise, even though the journal was originally founded in 1967.

[b] Early *Minerva* articles did not feature an abstract.

All collection ended in October 2019.

first identified with the journals' ISSN, using Crossref API, and HTML code from Sage web pages associated with articles was collected. We developed a parser to extract titles and abstracts for all articles[2]. Additionally, we used the Web of Science to collect reference sections from *SSS* and *SCI* to identify the citation pattern between the two journals. We employed the Python "Gensim" package to lemmatize word forms, remove stop-words, and classify part-of-speech. Table 1 shows the number of DOIs and articles included in our analyses from each journal. Additionally, Figure 1 displays the number of articles for each year.

3. CHARACTERIZING THE DIVIDE

With lemmatized texts from each article, we counted Boolean term frequency for each word to capture divergent word usages from the "two cultures" (Snow, 2012). We first computed the relative frequency for each word in both quantitative and qualitative articles. For example, 23.32% of articles from the quantitatively oriented journals used "impact" in the title or abstract, but only 5.85% of those from qualitative journals. Based on the relative frequency of words, we computed and arrayed the 10 most extreme ratios, separated by part of speech, in Figure 2. For example, the ratio of relative frequencies for "impact" from quantitative and qualitative journals, 3.99 (≈23.32/5.85), is shown in the first column, using the qualitative frequency as denominator. For words dominant in the qualitative side, numerators and denominators were switched. Note that we only considered words appearing more than 5% from both sides for Nouns, Verbs, Adjectives, and 3% for Adverbs or Prepositions for Figure 2. We provide an extended version with a 1% threshold for each part of speech in Figure A1 in the Appendix, and we will refer to word ratios as we discuss the divergence, even if they do not appear in the figures.

To capture nuanced semantic differences in language usage between the two cultures, we constructed two word embedding spaces using the titles and abstracts from each. Recent work in computational linguistics and natural language processing have powerfully represented entire systems of

[2] In some cases abstracts could not be extracted because they did not exist: DOIs sometimes indexed book reviews, editorial statements, short discussions, or obituaries, which we excluded from further analysis.

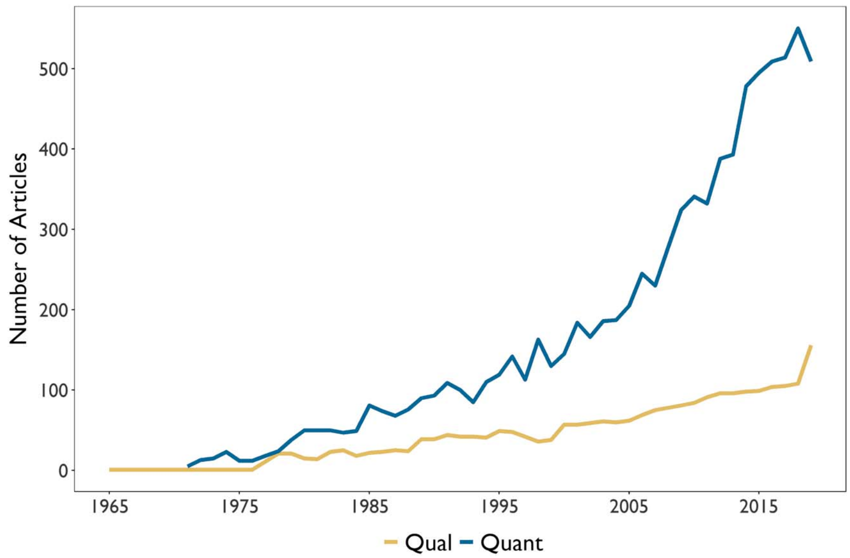

Figure 1. Aggregated number of articles for each category by year.

meaning by embedding words and sentences as vectors in dense, continuous, high-dimensional spaces (Le & Mikolov, 2014; Mikolov, Sutskever, et al., 2013; Pennington, Socher, & Manning, 2014). These vector space models, known collectively as *word embeddings*, have attracted widespread interest among computer scientists (Bolukbasi, Chang, et al., 2016; Levy and Goldberg, 2014), computational linguists (Garg, Schiebinger, et al., 2018; Hamilton, Leskovec, & Jurafsky, 2016), and social and behavioral scientists (Caliskan, Bryson, & Narayanan, 2017; Kozlowski, Taddy, & Evans, 2019) due to their ability to capture and represent complex semantic relations—including stereotypes, prejudice, and cultural association—inscribed within a discursive culture as present in a corpus of text.

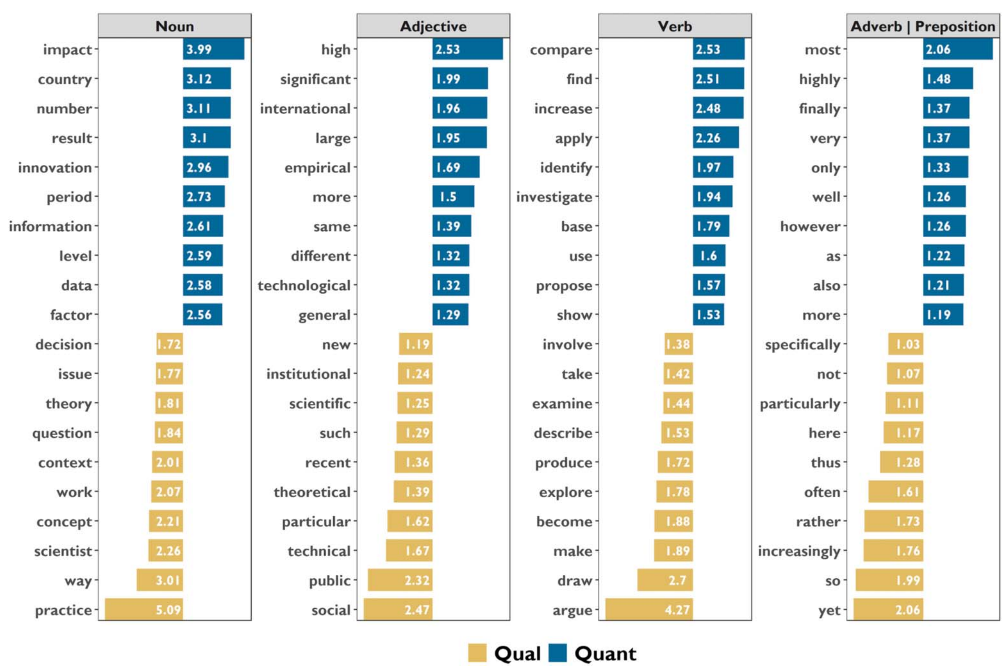

Figure 2. Ratio between probabilities of given tokens observed in abstracts and titles.

In a word embedding model, each word is represented as a vector in shared vector space with words sharing the similar surrounding words positioned nearby in the space. If *science* and *scientist* both appear near the word *laboratory*, then the vectors $\overrightarrow{science}$ and $\overrightarrow{scientist}$ will be located near each other in the embedding, even if they never appear together in the text. We use Google's word2vec (Mikolov et al., 2013), the most widely used word embedding algorithm, to construct two independent embeddings inscribing the cultures of quantitative and qualitative science studies. Each deployed the continuous-bag-of-words (CBOW) algorithm, used 50 dimensions, had a word window size of seven, and required each modeled word to appear in each corpus at least five times. Finally, to achieve and present robust results, we averaged projections from 100 trained models as word2vec incorporates a stochastic element that causes weight matrices to differ slightly with the same hyperparameters and corpus.

Within quantitative and qualitative epistemological cultures, we sought to identify the evaluative dimension along which those cultures array ideas and approaches as better or worse (Osgood, 1964; Osgood, Suci, & Tannenbaum, 1964). Word2vec initially received substantial attention based on its capacity to solve analogy problems, such as "man is to woman as king is to _____" (Mikolov et al., 2013). This can be solved by performing $\overrightarrow{king} + \overrightarrow{woman} - \overrightarrow{man}$, which will return a vector closest to the vector $\overrightarrow{queen}$ on a sufficient embedding space. This suggests that $\overrightarrow{woman} - \overrightarrow{man}$ inscribes a gender vector on which $\overrightarrow{queen}$ projects positively and $\overrightarrow{king}$ projects negatively. Building on this capacity, Kozlowski et al. (2019) proposed a method of constructing cultural dimensions such as class by taking the arithmetic mean of word vectors representing class antonyms (e.g., $\overrightarrow{rich}$, $\overrightarrow{affluent}$ and $\overrightarrow{poor}$, $\overrightarrow{impoverished}$). This approach has been widely validated and adapted (Ahn, 2020; An, Kwak, & Ahn, 2018; Bodell, Arvidsson, & Magnusson, 2019), and we employ it here to construct and compare quantitative and qualitative analysts' evaluation of different phenomena, concepts, and approaches to the study of science.

We anchored the evaluative dimensions in quantitative and qualitative science studies using the following word pairs: *good-bad*, *better-worse*, *right-wrong*, *satisfactory-unsatisfactory*, *positive-negative*, *sufficient-insufficient*, *effective-ineffective*, *excellent-failed*, *success-failure*. For example, to compute the association between *theory* in the quantitative worldview, we computed its orthogonal projection through calculation of the cosine similarity between the normalized word vector $\overrightarrow{theory}$ and the vector calculated by summing up $(\overrightarrow{good} - \overrightarrow{bad}) + (\overrightarrow{better} - \overrightarrow{worse}) + \ldots + (\overrightarrow{success} - \overrightarrow{failure})$ from the trained word2vec model based on the corpus from *SCI* and *RP*. A resulting value can fall between 1.0, signifying extreme positive evaluation, and −1.0, suggesting extreme negative evaluation. The value for *theory* on this evaluative projection in the quantitative embedding is −.17. We repeated the same procedures with the model based on the corpus from *SSS*, *STHV*, and *Minerva* to examine how the same words project in the qualitative world view, where *theory* projects much more positively at .04.

4. HOW QUANTITATIVE AND QUALITATIVE WORLD VIEWS DIVERGE

We first explore the relationship between quantitative and qualitative science studies by interpreting the nouns and adjectives they disproportionately use to capture the ontology of their worldviews—what is real in science to them. Then we investigate the verbs, adverbs, and prepositions they deploy to capture their investigative epistemology as they study science—reflecting how the two worlds differ in how they know what they know (Cetina, 2009). Finally, we explore the projections of words associated with divergent ontologies and epistemologies to examine how ontological and epistemological salience relates to evaluation. Of course, these three investigations are not mutually exclusive, so our mapping and projection of parts of speech involves some redundancy; but we believe that this allows us to understand, describe and reinforce what they know and what they value.

4.1. Qualitative Ontology

The first two columns of Figure 2 and Figure A1 reveal two contrasting ontologies or worldviews: They answer the question of what exists and is differentially worthy of consideration within the two cultures of science studies. Consider how the words *practice* (5.09), *scientist* (2.26), *work* (2.07), *actor* (3.57), *body* (4.50), and *material* (3.89) are hundreds of percent more likely to appear in qualitative than quantitative science studies. *Practice* appears 5.09 more times—509% more—in qualitative journals, reflecting that the *setting* (2.76) or *context* (2.01) of research and medical practice, such as the *laboratory* (3.62), have been a major context of qualitative studies, but remain invisible to quantitative researchers who only have access to statistics derived from the public record. For example, the *practices* of theoretical physicists have been studied by ethnographic observation (Merz & Cetina, 1997) and analysts have recently examined the shift in neuroscientific *practice* surrounding the advent of neuroinformatics (Beaulieu, 2001). Qualitative researchers have also lavished attention on cognitive practices and states deeply bound up with scientific and medical work, including *uncertainty* (2.73), *expectation* (2.75), *decision* (1.72), and *risk* (3.85). These practices are qualified with adjectives such as *experimental* (1.83), *professional* (3.98), *(bio) medical* (2.67), and *clinical* (2.58), suggesting a commitment to *institutional* (1.24) aspects of techno-scientific work that remain nevertheless *particular* (1.62) to distinctive domains of practice.

Qualitative science studies have also focused on the qualitative character or *form* (3.12) of *persons*[3] (3.05) inside scientific and medical contexts, such as *experts* (3.03). Beyond fixed characteristics, they pay special attention to processes—the *way* (3.01) events unfold—such as *transformation* (3.30)—and whether such processes are *social* (2.47), *cultural* (6.22), *conventional* (1.80), *political* (6.18), or *environmental* (2.76). They also consider large and amorphous social classes such as *publics* (3.21) and *society* (3.89) that exert influence within scientific settings, or forces that resist quantitative identification, such as *power* (2.80), *regulation* (3.85), and *governance* (3.20).

Qualitative analysts also undertake research in ways distinctive from those publishing quantitative journals. They *interview* (5.53) *participants* (3.57) and invoke *sociology* (5.22). They raise *issues* (1.77), *questions* (1.84), and *concerns* (3.24), make *arguments* (3.09), and engage in *debate* (5.32). They cultivate *notions* (4.72) and *constructions* (5.25) that mature into *concepts* (2.21) and constitute *theory* (1.81). Qualitative researchers also manifest different rhetorics of argumentation from their quantitative fellows. They forward positions *central* (2.08), *key* (1.23), *crucial* (1.78), and *fundamental* (1.83), but they have affection for perspectives at the periphery that are *critical* (1.89), *alternative* (1.78), and *distinct* (1.88).

4.2. Quantitative Ontology

The quantitative landscape of scientific subjects and objects appears dramatically different. Rather than examining the qualities of scientific practice, they focus on *quantitative* (2.72) and *economic* (2.87) outputs of scientific organizations including *performance* (3.98), *impact* (3.99), *investment* (3.15), and *innovation* (2.96). Rather than materiality and practice, they attend to legal and

[3] We do not distinguish between plural and singular words in our analysis, which considered lemmatized words, but vary them in text to facilitate exposition.

corporate entities including *firms* (12.14) and *countries* (3.12), which are identifiable from published scientific metadata, and describe things with respect to the boundaries of these legal objects—as *foreign* (3.77), *regional* (3.57), or *international* (1.96). They discuss corporate scientific entities in terms of whether they are *better* (2.56) or worse, *productive* (2.49) or inert, and *collaborative* (2.34) or independent. They also consider measurable traces of the publication process, such as *publication* (7.08), *journal* (6.74), and *citation* (12.36) using the *web* (7.84) of science, but also *patents* (10.83). Quantitative science studies researchers view these facets of scientific action as *measures* (4.45) or *indicators* (12.12) of higher level phenomena such as *collaboration* (2.89). Moreover, they describe them not in the context of a case or example, but rather as a *sample* (3.99) of an underlying *distribution* (3.26).

Quantitative science studies researchers see their data in terms of *number* (3.11), analyzed by some *method* (1.77) or *methodology* (2.79) to evaluate a *hypothesis* (3.08). This allows them to discover a *result* (3.10) in the form of an *increase* (5.53), *trend* (4.83), *level* (2.59), *share* (2.99), or *combination* (2.75).

Another ontological divergence is found in the way the two cultures predicate their subjects of study. Referring to a statistical model, quantitative analyses tend to report *significant* (1.99) findings. Quantitative science studies (obviously) evaluate these subjects quantitatively, in terms of whether they are the *same* (1.39) or *different* (1.32), represent a *simple* (2.11) or *full* (2.15) model specification, or score *high* (2.53) or *low* (3.81), *large* (1.95) or *small* (2.33), *more* (1.50) or *less* (1.25), and *positive* (4.11) or *negative* (2.73) on some metric. Discursively, quantitative articles hierarchically rank their findings, easily summarized by *main* (2.59), *overall* (3.24), and *general* (1.29) points. They qualify these findings as being *consistent* (2.35) or *relative* (2.88) to one another, and isolate outcomes *due* (2.22) to the same underlying causes.

In striking contrast, quantitative versus qualitative science studies tend to draw upon *economics* (2.87) versus *sociology* (5.22); focus on objects *international* (1.96) versus *institutional* (1.24), *general* (1.29) versus *particular* (1.62), *empirical* (1.69) versus *theoretical* (1.39), and *simple* (2.11) versus *complex* (1.51); and finally frame their arguments in terms of *hypotheses* (3.08) versus *questions* (1.84) and *findings* (2.18) versus *issues* (1.77).

The relative prevalence of verbs and adverbs in the two cultures of science studies provides powerful insight to the research behaviors that distinguish qualitative versus quantitative epistemologies.

4.3. Qualitative Epistemology

Qualitative researchers disproportionately *explore* (1.78) scientific phenomena, *looking* (2.06) for and *seeking* (2.71) to *recognize* (2.15) and *understand* (2.65) insights they subsequently *describe* (1.53) to their audiences. They articulate their tasks and those of the agents they observe with a physicality uncharacteristic of quantitative work. They *engage* (2.60) and *call* (2.13), *bring* (2.86) and *take* (1.42), *come* (2.00) and *go* (2.07), and *move* (2.81) and *turn* (2.64). This may be unsurprising as qualitative research methods inherently involve the body as both instrument and object of study. Qualitative researchers discursively unfold their exposition through active, even muscular, rhetoric, suggesting explicit acts of construction. They *claim* (4.63), *argue* (4.27), *challenge* (3.10), *maintain* (2.43), and *emphasize* (2.48) ideas in building up a system that *engages* (2.60) other concepts, arguments, and interlocutors.

Insofar as social constructivism implies that human action is socially situated and knowledge is constructed through interaction, qualitative science studies scholars perform this approach (Collins, 1981). They *draw* (2.70), *shape* (4.23), *embed* (2.34), *constitute* (3.06), and explicitly *construct* (2.40) theories that not only reflect their propensity to *view* (2.10), *see* (2.89), and *look* (2.06) at scientific phenomena in situated *context* (2.01) but also produce knowledge in the situated context of the article form itself. As such, they are not certain of their claims, but *attempt* (1.91) them.

Moreover, their mode of research activities resists strict quantification. Observations are made *closely* (1.63), and the phenomena they observe occur *always* (1.49), *often* (1.61), *usually* (1.37), and *largely* (1.51). Spatially and temporally, things happen *here* (1.17), *now* (2.02), *currently* (1.30), or *increasingly* (1.76). Qualitative students of science muse on the cosmic space of possibilities by considering things that could even occur *potentially* (1.72). This reflects how subjects of qualitative research are themselves engaged in complex transformations that resist quantification: They *arise* (2.69), *become* (1.88), and *change* (1.89), but then *continue* (2.10) in the paths they *begin* (2.47). Violating the canonical iid assumption behind most econometric models (that cases are identically and independently distributed), most things happen *together* (2.01). Defying the causal ordering of events, many of those things occur *simultaneously* (2.68). This complexity requires contingent argumentation, and the creation and subtle resolution of paradox, reflected by intensive use of prepositions *rather* (1.73), *so* (1.99), *yet* (2.06), and *instead* (1.86).

4.4. Quantitative Epistemology

Quantitative researchers follow the central dogma of empiricism: They *observe* (2.15), *collect* (2.21) *information* (2.61), and *measure* (4.63) *data* (2.58), then *apply* (2.26), *use* (1.60), or *utilize* (1.72) methods to *identify* (1.97) and *find* (2.51) things about the world. They do not explore, they *investigate* (1.94). Rather than constructing new theories, they *assess* (2.55) and *compare* (2.53) existing ones, *proposing* (1.57) hypotheses, then *evaluating* (2.54), and *testing* (2.83) them. This poises quantitative students of science to use their validated *models* (1.35) to *predict* (2.37) the future and *confirm* (5.22) what they *expect* (2.72).

A key epistemological move in the quantitative *habitus* is to identify inexpensive quantities that *indicate* (3.06) concepts of theoretical interest, and then *analyze* (1.65) the structure of indicators assuming a conserved *map* (2.28) or homomorphism with the pattern of underlying concepts. Unlike qualitative analysts of science, they assume an attitude of objectivity, which leads to greater certainty: They do not *question*, they *determine* (2.45). They do not *construct*, they *obtain* (4.34).

Insofar as quantitative science studies research *innovation* (2.96), *productivity* (5.13), *performance* (3.98) and *outputs* (10.78), they naturally observe these states *increase* (2.48) *improve* (2.44), *enhance* (1.78) and *perform* (1.76). Their adverbs illuminate how quantitative science analysts do what they do. They make their arguments *empirically* (1.60) and their models reveal findings *significantly* (3.22). They also reveal quantities underlying their claims. They observe phenomena *more* (1.18), *frequently* (1.74), *highly* (1.48), and *recently* (1.37). Quantification enables them to discuss scientific and technological agents that are *most* (2.06) and *very* (1.37). The structure of their statistical models allow them to *compare* (2.53) and account for things *relatively* (2.28). Quantitative analysts of science make claims declaratively, like a high-level programming language. Their findings may be hierarchically summarized *mainly* (2.18) or *generally* (1.26), and listed serially, *furthermore* (2.74), *also* (1.21), and *therefore* (1.56), concluding *finally* (1.37) in contrast with the complex contingencies in qualitative exposition. To summarize, acolytes of empirical epistemologies *test* (2.83), while interpretive explorers *argue* (4.27). Quantitative analysts objectively *observe* (2.15) and *find* (2.51), while qualitative scholars intuitively *see* (2.89) and *understand* (2.65). These differences in creative agency burrow down to the atoms of quantitative and qualitative analysis: Quantitative findings *consist* (1.73) of discovered facts while qualitative ones are *constituted* (3.06) by them.

5. EVALUATION

To capture how quantitative and qualitative science studies evaluate their own most common ontological observations and epistemological moves, but also those of the other, we embedded all of their words into two separate high-dimensional Euclidean spaces, enforcing internal semantic consistency to quantitative and qualitative articles, respectively. We view quantitative and qualitative embedding models as representations of the worldviews distinct to the two cultures. Then we induced an evaluative dimension that distinguished how those worldviews judged research approaches (verbs and adverbs), subjects of study (nouns and adjectives), and discursive moves (prepositions, conjunctions, etc.) We initially selected 10 words prior to computation to illustrate their difference between qualitative and quantitative worldviews, as shown in Figure 3. The evaluative projection of these words within quantitative and qualitative worldviews manifest strong negative correlations—Spearman correlation of −.71 and Pearson correlation of −.50.—suggesting that quantitative and qualitative science studies reflect not independent but opposing epistemological worlds. They know different things, but they also value that knowledge in opposing ways. Moreover, when we take all the words from Figure 2 and project them on the evaluative dimension in each of the two worlds, we continue to find strong negative correlations of −.53 (Spearman), and −.50 (Pearson)—see Figure A2 in the Appendix.

This relationship between frequency of use and positive evaluation suggests a strong relationship between what we know and what we prefer. By observing an equally potent relationship between negative evaluation and infrequency of use in one's worldview, we see that researchers not only study but also laud what they believe can be studied and denigrate what they believe cannot. Moreover, it appears that what one culture of science studies examines infrequently—but their neighbor examines intensively—are things they believe neither can nor should be studied. This represents bad work. Our findings delineate not only ontological and epistemological but also moral boundaries between quantitative and qualitative studies of science, where what can be studied maps onto what should be studied.

The citation pattern between journals reflects this division. Figure 4 displays the citation pattern between *SCI* and *SSS* from 2000 to 2019 based on a rolling average of the proportion of papers in each journal citing the other. For example, the value for 2000 was calculated as the nonweighted mean of the citing ratio from 1996 to 2000. The figure demonstrates persistent division. But the division is curiously asymmetric. *SCI* papers have been more likely to cite *SSS* papers than the

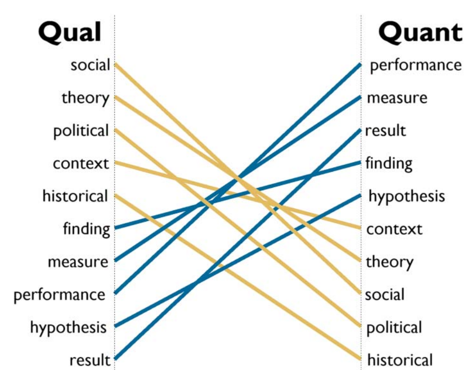

Figure 3. Projection of words onto the evaluative ("good-bad") dimension in each culture.

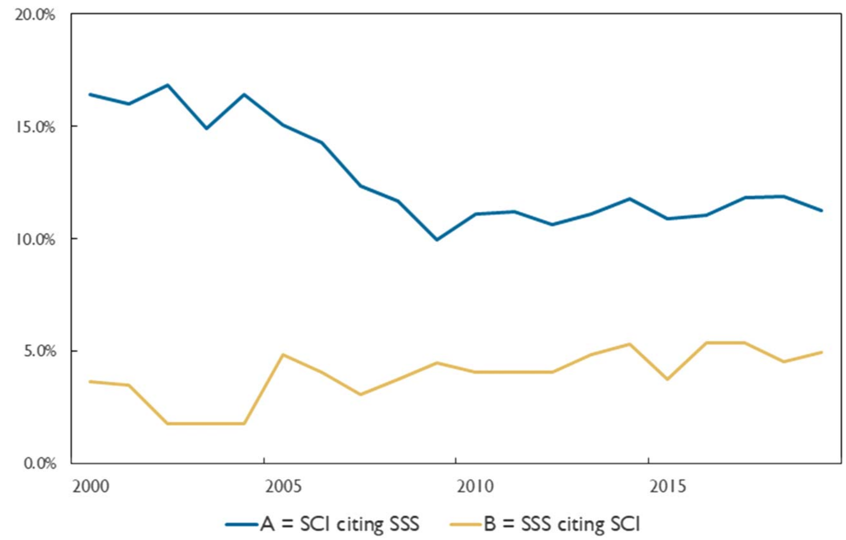

Figure 4. Five-year rolling average of the proportion of papers with cross-citation between *SCI* and *SSS* from 2000 to 2019.

converse, suggesting that quantitative science studies sometimes employ qualitative studies to propose theories and provide framing. But even this has shrunk in recent years.

6. BREAKING THE BARRIER

The meager, asymmetric collaboration we observe between qualitative and quantitative science studies—qualitative research discovers things worth counting, then quantitative research counts them—suggests a fixed division of labor that systematically limits the one-way insights that can pass between the cultures. Combined with the limited proclivity of quantitative sciences studies to engage in theory-building, or qualitative investigations to evaluate hypotheses, the potential for advance in either has become severely constrained. But one can imagine a new treaty between them for the benefit of science about science (Fortunato, Bergstrom, et al., 2018).

Patterns above the level of ethnographic visibility, detectable only through surveying large-scale quantitative patterns, could lead directly to new discoveries in the science of science. If quantitative science studies adopted a complex systems approach that not only tested preimagined hypotheses but also documented emergent phenomena, then they might have insights to intrigue and feed qualitative investigation that had been ignored, only anchored and obvious above the level of observable scientific practice. Lee and Martin (2015) argue for this possibility of quantitative "cartography—the construction of question-independent, though theoretically organized, reductions of information to make possible the answering of many questions."

Regularities in quantitative traces of science demand complex causal investigation, which in turn require qualitative assessment (Small, 2013; Tavory & Timmermans, 2013) in that it involves new and unanticipated factors, which may be visible through the instrument of the body even though never previously collected. As a result, qualitative research facilitates the detection of complex configurations of delicate signals, which could unearth chains of causal influence, or test and compare quantitatively derived inferences. For example, in a recent paper we published (Wu, Wang, & Evans, 2019), we identified a strong, negative relationship between the size of teams producing science and technology, and their likelihood to disrupt the frontier in their domains, controlling for authors and outputs visible across 65 million teams. We did not publish a causal account for *how* this occurs, because the complex mixture of cognitive and social phenomena underlying the effect, ranging from collaborative inhibition (Barber, Harris, & Rajaram, 2015) to risk aversion (Christensen, 2013) to transaction costs (Williamson, 1985) to some other unimagined force, involve factors that have not been, and may never be, measured across the tens of millions of teams we examined quantitatively.

In recent years, large-scale digital repositories of complex, qualitative scientific artifacts have become widely available through online services such as Figshare, the Open Syllabus project, GitHub, and customized collections drawn from the web or through recording and digitization efforts. The supply of these digital artifacts, including preprints, proposals, slide presentations, meeting recordings, and online conversations, has driven demand for powerful machine learning tools, especially deep neural networks (Manning, 2015), that can encode them into quantitative data. But these artifacts are fundamentally high-dimensional, containing many overlapping qualities—each of which could potentially be enumerated. What to quantify? If we reduce these artifacts into quantities previously theorized, we ignore the vast new views of science uniquely available to us. Emergent patterns in complex data demand qualitative sensibilities to determine a research focus and how to interpret, theorize, and qualify insights (Evans & Foster, 2019; Nelson, 2017). But these skills have been underexercised and undervalued among quantitative researchers. Bringing together qualitative sensibility with quantitative literacy and computational skill will require overcoming ontological, epistemological, and moral divides regarding what is real, what is knowable, and what is good.

In an era of small data on science, quantitative approaches maximized our insight by testing strong theories, buttressed by a myriad assumptions. In an era of big data, however, we maximize discovery by reducing our assumptions, weakening our theories and growing new ones (Lee & Martin, 2015). With big data, we can inductively discover grounded theory on some data (Hannigan, Haans, et al., 2019; Nelson, 2017), then quantitatively test it on other data. But doing this requires removing the misplaced moral taint associated with "data mining" embedded within the contemporary quantitative culture of science studies. The importance of fixed, preset hypotheses made sense in an era of small data, but does not in one of large data. We believe that new computational methods and digital data could weaken the demarcation between quantitative and qualitative science studies, if we can overcome the evaluative commitments that separate them. New computational methods and digital data could also broach another divide.

Quantitative science studies, as published in outlets such as *RP*, often address the agents of science policy, such as public funders or journal editors. By identifying bias, and pathways to higher performance critical to the mission of these policy-makers, such studies offer direct insights that could improve the institutions of science. Qualitative science studies, by contrast, often reflect on the implications of closely examined scientific institutions and policies. Articles in serials such as *Minerva* and *Nautilus* raise broader questions about how and why science is as it is, and what alternatives might exist. But neither tends to speak to the scientists they study. Historically, the philosophy of science dealt with the making and defending of scientific claims (Quine, 1951) and the public moves scientists undertake, which drive both scientific influence and impact. In the mid-20th century, philosophers of physics would publish in physics journals, arguing over the legitimate interpretation of observations and experiments.

We argue that discursive acts of claims-making are critical behaviors in science, and that with computational methods and large-scale samples of this activity, science studies can begin to address questions that directly address scientists, their efforts, and the scientific and reputational consequences that follow. A reformulated treaty between quantitative and qualitative studies, forged by computational methods and big scientific text data could portend a renaissance of engagement with scientists about what they do—forward and defend scientific claims—at the level at which they do it. Deeper collaboration between qualitative and quantitative studies of science would span scales of analysis, connecting scientific action to science policy along one dimension, and civic action and participation on the other (Durant, Evans, & Thomas, 1989), potentially driving new discoveries and new relevance.

We believe that our own case, contrasting the relationship between qualitative and quantitative science studies, demonstrates how computation can disturb the demarcation between quantity and quality. The enumeration of word quantities requires in-depth interpretation, which involves theorizing about distinct qualities. While exploiting computational representations emerging from our models built from qualitative and quantitative research articles, we had to consider nuances of meaning about words in context to make sense of their projections on the evaluative dimension within each of the two cultures.

We note that our observed divisions between quantitative and qualitative science studies reflect broader patterns distinguishing quantitative and qualitative turns across social and natural sciences. Qualitative and quantitative science studies were much more likely to reference sociology and economics, respectively. These index homologous differences in qualitative and quantitative analysis of governments, companies, markets, schools, and other domains of social life. Qualitative approaches, more prevalent in sociology, anthropology, and history, examine transformations, conjunctions and other complex social processes resisting quantification. Quantitative methods, central to economics, and increasingly political science, statistics, and computer and information science, focus more frequently on defensively establishing increases, decreases, corelations, and causes. These patterns are not limited to the social sciences. Peter Galison has documented the mid-20th-century conflict between qualitative, image-based analyses in physics using exposed film and cloud chambers to visually trace particles, versus logical tests of particle presence with spark chambers (Galison, 1997). In Evelyn Fox Keller's *A Feeling for the Organism*, she details a stark distinction between Barbara McClintock's qualitative, intuitive understanding of maize and genetics, contrasting it with formal models and small-scale controlled experiments of dismissive colleagues like Sewall Wright and Joshua Lederberg (Keller, 1984)[4]. In many historical cases, however, the merger of qualitative insights and quantitative formalisms and large-scale measurement have been associated with breakthrough advance, as in the construction of bubble chambers and the formal characterization of genetic operons and transposons.

Auguste Comte argued that the practice of science in action should also be a subject of an empirical investigation (Comte, 1988), but only a new pact between qualitative and quantitative investigation can approach this aspiration. It leads us to sympathize with Feyerabend: "The idea of a fixed method, or a fixed theory of rationality, rests on too naïve a view of man and his social surroundings. To those who look at the rich material provided by history, and who are not intent on impoverishing it in order to please their lower instincts, their craving for intellectual security in the form of clarity, precision, 'objectivity,' 'truth,' it will become clear that there is one principle … anything goes" (Feyerabend, 1975, p. 18). Feyerabend's indeterminacy goes beyond current evidence—stable, conservative institutions of science have ensured for centuries that not *anything* goes—but he defensibly highlights that advance is typically associated with a violation of existing methodological boundaries. With a wealth of new digital data on science and computational methods to explore it, we could squander this opportunity to understand science in new ways,

[4] Quantification is not always ascendant. Several contemporary biological publications are moving toward visual abstracts, while biology-related equations may reduce the spread of their ideas in broader biological discourse (Fawcett & Higginson, 2012).

or we could move against method—transgressing the boundary of epistemic cultures for a deeper collective understanding of the scientific enterprise and our role within it, portending a Renaissance in science studies.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

This research was supported by a grant from AFOSR FA9550-19-1-0354.

DATA AVAILABILITY

Data and code for analysis available at: https://github.com/KnowledgeLab/AgainstMethod/.

REFERENCES

Ahn, Y. Y. (2020). FrameAxis: Characterizing framing bias and intensity with word embedding. *arXiv:2002.08608v2*.

An, J., Kwak, H., & Ahn, Y. Y. (2018). SemAxis: A lightweight framework to characterize domain-specific word semantics beyond sentiment. *arXiv Preprint arXiv:1806.05521*. http://arxiv.org/abs/1806.05521

Barber, S. J., Harris, C. B., & Rajaram, S. (2015). Why two heads apart are better than two heads together: Multiple mechanisms underlie the collaborative inhibition effect in memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *41*(2), 559–566.

Beaulieu, A. (2001). Voxels in the brain: Neuroscience, informatics and changing notions of objectivity. *Social Studies of Science*, *31*(5), 635–680.

Bodell, M. H., Arvidsson, M., & Magnusson, M. (2019). Interpretable word embeddings via informative priors. *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/1909.01459

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, *29*, 4349–4357.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186.

Cetina, K. K. (2009). *Epistemic cultures: How the sciences make knowledge*. Harvard: Harvard University Press.

Christensen, C. M. (2013). *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard: Harvard Business Review Press.

Collins, H. M. (1981). Stages in the empirical programme of relativism. *Social Studies of Science*, *11*(1), 3–10.

Comte, A. (1988). *Introduction to positive philosophy*. Cambridge, MA: Hackett Publishing.

Durant, J. R., Evans, G. A., & Thomas, G. P. (1989). The public understanding of science. *Nature*, *340*(6228), 11–14.

Evans, J., & Foster, J. G. (2019). Computation and the sociological imagination. *Contexts*, *18*(4), 10–15.

Fawcett, T. W., & Higginson, A. D. (2012). Heavy use of equations impedes communication among biologists. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(29), 11735–11739.

Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. Humanities Press.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., … Barabási, A.-L. (2018). Science of science. *Science*, *359*(6379). https://doi.org/10.1126/science.aao0185

Galison, P. (1997). *Image and logic: A material culture of microphysics*. Chicago: University of Chicago Press.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), E3635–44.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/1605.09096

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., … Jennings, P. D. (2019). Topic modeling in management research: rendering new theory from textual data. *Academy of Management Annals*, *13*(2). https://doi.org/10.5465/annals.2017.0099

Keller, E. F. (1984). *A feeling for the organism, 10th Anniversary edition: The life and work of Barbara McClintock*. Basingstoke: Macmillan.

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949.

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *International Conference on Machine Learning*. http://www.jmlr.org/proceedings/papers/v32/le14.pdf

Lee, M., & Martin, J. L. (2015). Surfeit and surface. *Big Data & Society*. https://doi.org/10.1177/2053951715604334

Levy, O., & Goldberg, Y. (2014). Neural word embedding as implicit matrix factorization. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.) *Advances in Neural Information Processing Systems 27* (pp. 2177–2185). Red Hook, NY: Curran Associates, Inc.

Leydesdorff, L., & Van Den Besselaar, P. (1997). scientometrics and communication theory: Towards theoretically informed indicators. *Scientometrics*, *38*(1), 155–174.

Manning, C. D. (2015). Computational linguistics and deep learning. *Computational Linguistics*, *41*(4), 701–707.

Merz, M., & Cetina, K. K. (1997). Deconstruction in a 'thinking' science: Theoretical physicists at work. *Social Studies of Science*, *27*(1), 73–111.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.) *Advances in Neural Information Processing Systems 26* (pp. 3111–3119). Red Hook, NY: Curran Associates, Inc.

Milojević S., Sugimoto, C. R., Larivière, V., Thelwall, M., & Ding, Y. (2014). The role of handbooks in knowledge creation and diffusion: A case of science and technology studies. *Journal of Informetrics*, *8*(3), 693–709.

Nelson, L. K. (2017). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3–42. https://doi.org/10.1177/0049124117729703

Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, *66*(3), 171–200.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1964). *The Measurement of Meaning*. University of Illinois Press.

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* Doha (pp. 1532–1543). Association for Computational Linguistics.

Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *Philosophical Review*, *60*(1), 20–43.

Small, M. L. (2013). Causal thinking and ethnographic research. *American Journal of Sociology*, *119*(3), 597–601.

Snow, C. P. (1959). *The two cultures and the scientific revolution*. Cambridge: Cambridge University Press. http://sciencepolicy.colorado.edu/students/envs_5110/snow_1959.pdf

Snow, C. P. (2012). *The two cultures*. Cambridge: Cambridge University Press.

Tavory, I., & Timmermans, S. (2013). A pragmatist approach to causality in ethnography. *American Journal of Sociology*, *119*(3), 682–714.

Teplitskiy, M., Acuna, D., Elamrani-Raoult, A., Körding, K., & Evans, J. (2018). The sociology of scientific validity: How professional networks shape judgement in peer review. *Research Policy*, *47*(9), 1825–1841.

Williamson, O. E. (1985). *The economic institutions of capitalism*. New York, NY: Free Press.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*, 378–382. https://doi.org/10.1038/s41586-019-0941-9

APPENDIX

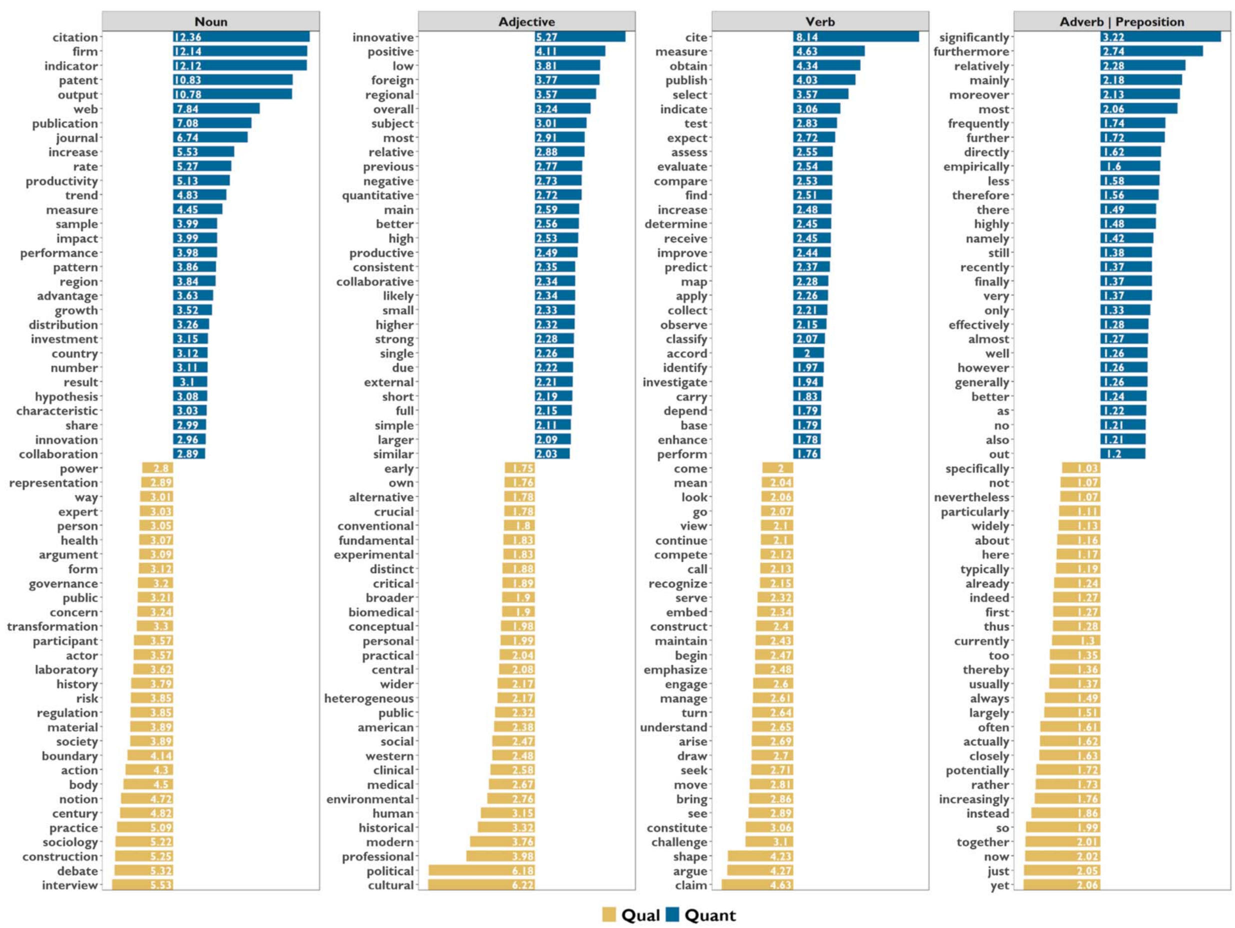

Figure A1. Extended comparison of prevalent words using a 1% threshold: Words are reported if they appear more than 1% of the time from the more frequent corpus in the less frequent one.

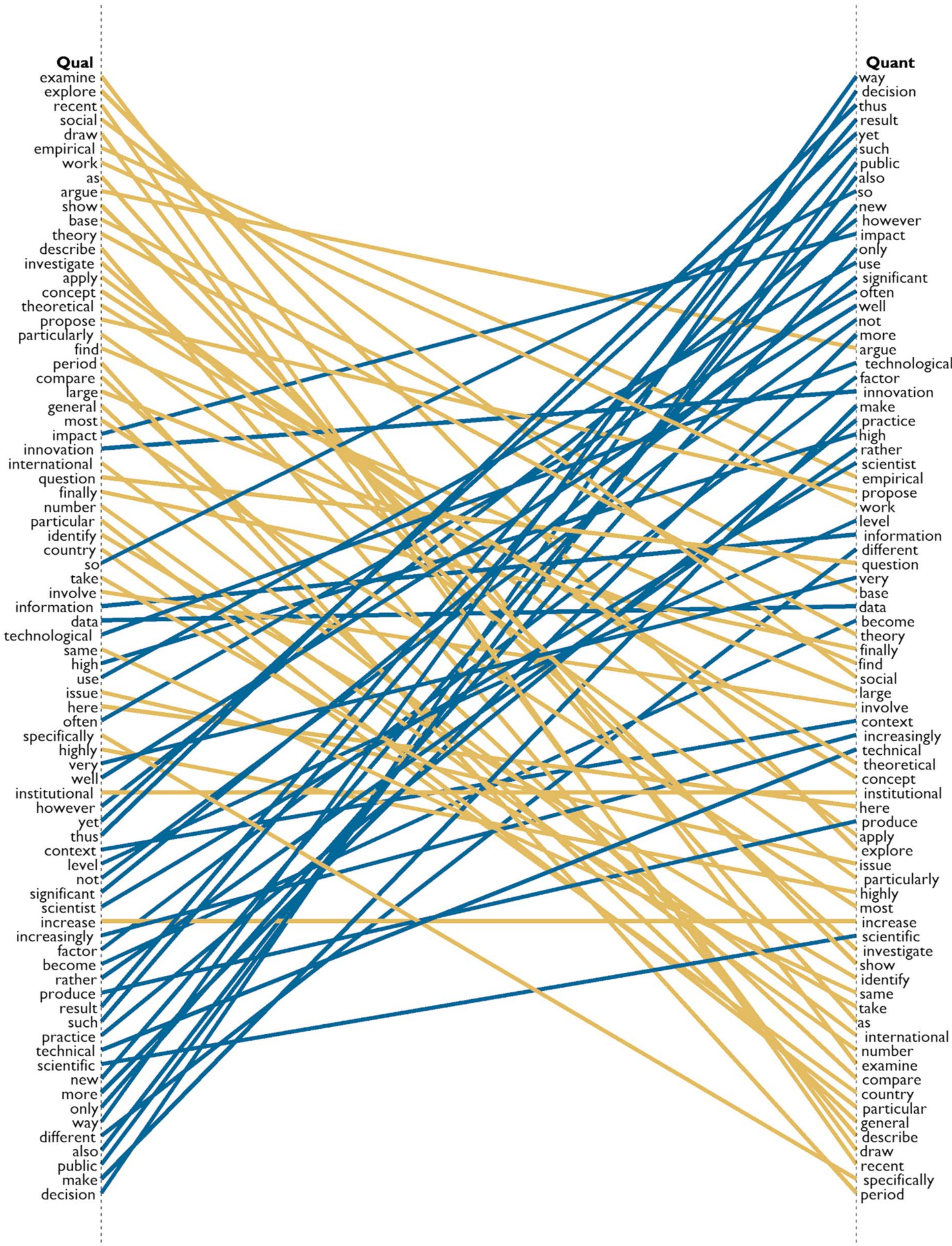

Figure A2. Rank order of all words in Figure 2, projected onto the evaluative ("good"/"bad") dimensions inscribed by qualitative and quantitative science studies articles.