Toward predicting research proposal success

3 authors:

Kevin Boyack
SciTech Strategies, Inc.
122 PUBLICATIONS **4,525** CITATIONS

SEE PROFILE

Caleb Smith
University of Michigan
2 PUBLICATIONS **4** CITATIONS

SEE PROFILE

Richard Klavans
SciTech Strategies, United States
77 PUBLICATIONS **2,878** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Call for Papers: Mapping Science: Resources, Tools, and Benchmarks (#FrontiersInRMA) View project

Project   Topic identification challenge View project

Toward Predicting Research Proposal Success

Kevin W. Boyack[1], Caleb Smith[2] & Richard Klavans[3]

[1] *kboyack@mapofscience.com*
SciTech Strategies, Inc., Albuquerque, NM (USA)

[2] *calebs@med.umich.edu*
University of Michigan Medical School, Ann Arbor, MI (USA)

[3] *rklavans@mapofscience.com*
SciTech Strategies, Inc., Wayne, PA (USA)

Abstract: Citation analysis and discourse analysis of 369 R01 NIH proposals are used to discover possible predictors of proposal success. We focused on two issues: the Matthew effect in science – Merton's claim that eminent scientists have an inherent advantage in the competition for funds – and quality of writing or clarity. Our results suggest that a clearly articulated proposal is more likely to be funded than a proposal with lower quality of discourse. We also find that proposal success is correlated with a high level of topical overlap between the proposal references and the applicant's prior publications. Implications associated with the analysis of proposal data are discussed.

Introduction

There are both theoretical and practical reasons for analyzing research proposal success and failure. From a theoretical perspective, it is important to realize that much of the recent progress in modeling how science evolves is based on the analyses of tens of millions of scientific articles. These articles, by their very nature, look backwards in time. They represent research supply (Sarewitz & Pielke Jr., 2007). In contrast, the millions of research proposals that are stored in the archives of universities and funding agencies (and are currently unavailable for large-scale analysis) each looked forward in time at the time of their submission. Research proposals are at the interface of supply and demand. They represent the supply of research in that they are documents regularly produced by researchers. They are also related to research demand in that they frame their proposed research in terms of relevant societal issues (such as disease burden, capability to respond to pandemics, threat of climate instability or economic opportunities associated with artificial intelligence). Significant progress in modeling the science of science should accrue if we build models from research proposals that are at the interface of supply and demand –and link these research proposals to our existing models of the scientific literature that reflect supply.

From a practical perspective, the average proposal by a scientist in the United States takes about 200 hours to prepare and costs about $20,000[1]. Current proposal success rates for proposals

---

[1] The average amount of time required to prepare a single proposal ranges from 170 hours (Von Hippel & Von Hippel, 2015) to 270 hours (Herbert, Barnett, Clarke, & Graves, 2013) of researcher or investigator

submitted to U.S. agencies hover around 20%. Mathematically, a scientist that writes a proposal for $100,000 is only covering the proposal writing costs! Understanding what features are associated with successful vs. unsuccessful proposals is particularly important for early stage faculty who may not have the time or institutional commitment to write three or four proposals that fail before they write one that succeeds. While it's true that early career researchers are in 'publish or perish' environment, it's also true that early career researchers will flourish if they can become rainmakers – if they can bring grant money into their institutions.

It is in this context that we conducted a small proof of principle study to specifically identify the features of research proposals that could be used to predict the likelihood of proposal success. Building on Swales (1986), we focus on two methodologies: citation analysis and discourse analysis. Citation analysis is used to test for the Matthew effect – that eminent scientists have an inherent advantage in the competition for funds. Discourse analysis is used to test for clarity under the hypothesis that proposals that are more clearly written are more likely to be accepted.

The structure of this article is as follows. We start with a review of the prior literature and highlight the gaps we will be addressing. The research design is then described. Our data sample consisted of 369 R01 proposals submitted to the U.S. National Institutes of Health (NIH) by the University of Michigan Medical School (UMMS) in 2010. We then present our findings, which are compelling. Consistent with prior work, there's only a small shred of evidence in support of the Matthew effect. This calls into question the tendency to emphasize research leadership in hiring and promotion. In contrast, proposal clarity was a surprisingly effective predictor of proposal success. We conclude by discussing how discourse analysis of research proposals might contribute to the theory of 'the science of science' and improve researcher productivity.

Prior Literature

Most of the peer reviewed empirical literature that looks at predicting proposal success focuses on the Matthew effect in science – the hypothesis that eminent scientists and/or eminent research institutions have an inherent advantage in the competition for funding. The Matthew effect in science was popularized by Merton (1968), and was based on original research done by Zuckerman (1967) on the tendency for Nobel Prize winners to be over-acknowledged. Building upon this phenomenon, Merton was correspondingly concerned with the increasing concentration of research funding amongst eminent research universities and specifically referred to this phenomenon as the Matthew effect in science.

> "When the Matthew effect is thus transformed into an idol of authority, it violates the norm of universalism embodied in the institution of science and curbs the advancement of knowledge. But next to nothing is known about the frequency with which these practices are adopted by the editors and referees of scientific journals and by other gatekeepers of science. This aspect of the workings of the institution of science remains largely a matter of anecdote and heavily motivated gossip (Merton, 1968, p. 7)".

---

time. This equates to roughly $8,500 to $13,400 USD in salary costs alone. Actual costs per proposal can be $20,000 on average when administrative overhead rates are included.

The first large scale study of the Matthew effect in the proposal review process was conducted 10 years later for the National Science Foundation (NSF). In this study, there was a small correlation between the awarding of grants and being an eminent scientist (Cole, Rubin, & Cole, 1978). For the 1200 applicants in this study over 10 program areas, 6% of the variance in the reviewer's rating was attributed to their past publication record. The authors considered this inconsequential because one of the three major criteria for a reviewer's assessment was the competence of the principal investigator. Additional analyses were done using alternative indicators such as prior awards by the applicant. Cole et al. correspondingly concluded that grant awards were not based on the eminence of the scientist or the eminence of the institution from which they came.

Subsequently, there have been a number of studies reporting similar findings. Several authors found that the productivity and impact of the applicant (before submission of a grant application) were higher for funded than unfunded applicants (Bornmann & Daniel, 2005, 2006; Cabezas-Clavijo, Robinson-Garcia, Escabias, & Jimenez-Contreras, 2013; Van den Besselaar & Leydesdorff, 2009). However, when funded applicants were compared with the best unfunded applicants using equal sample sizes, differences in performance went away (Bornmann, Leydesdorff, & van den Besselaar, 2010; Van den Besselaar & Leydesdorff, 2009). In these cases the differences between funded and unfunded applicants were entirely due to the tail of the unfunded distribution, which comprised at least 60% of the applicants. Self-selection appears to be an important factor. In some grant competitions, only the most qualified potential candidates or groups applied for funding (Enger & Castellacci, 2016; Van Leeuwen & Moed, 2012), and differences in performance between those ultimately funded and not funded could not be detected. There are also studies that have found that funded and unfunded applicants have similar pre-application productivity and impact (Hornbostel, Böhmer, Klingsporn, Neufeld, & Von Ins, 2009; Melin & Danell, 2006; Neufeld & Hornbostel, 2012; Neufeld, Huber, & Wegner, 2013; Saygitov, 2014). Overall, previous tests for the Matthew effect in science do not provide a compelling case that the rich are getting richer. Eminent scientists are only slightly more likely to be funded than less eminent scientists.

There are also studies that have looked at the Matthew effect from the opposite perspective, testing to see if those who are funded will become more eminent. On the one hand, studies comparing funding decisions with post-award performance have shown that funded applicants had higher productivity and impact than unfunded applicants (Bornmann, Wallon, & Ledin, 2008; Reinhart, 2009). However, these same studies have also shown high levels of type II errors (unfunded applicants with subsequent high performance), and have acknowledged that any performance differences found may have been explicitly due to the increased opportunity afforded by the funding (Van den Besselaar & Sandström, 2015). These studies, taken together, suggest that when the applicant pool is restricted to the most well-qualified or highest performing applicants, whether by peer review threshold or self-selection, it is very difficult to distinguish the best applicants from the second tier. Along the same lines, Jacob & Lefgren (2011) compared the outputs of funded and unfunded applicants for NIH R01 grants, showing that receipt of one of these grants had only a small relative effect on the research output (one additional paper over five years) of the marginal applicants.

There are related studies that focus on the predictive value of the peer review process. These studies exclude the unfunded proposals and shift the focus of analysis to project characteristics instead of applicant characteristics. The question at hand is whether the peer review scores (for the set of funded proposals) predict which projects will have higher productivity or impact. The results have been ambiguous. Gallo et al. (2014) showed moderate correlation between peer review scores and total time-adjusted citation output of funded projects for a small number of grants, while Li & Agha (2015) showed high positive correlations between NIH percentile scores and impact for a set of 130,000 grants. Although this seemed definitive at the time it was published, Fang et al. (2016) recently re-evaluated the data used by Li & Agha, restricting their analysis to the 102,740 projects with a percentile score of 20 or higher, and found no correlation between citation metrics and percentile scores. The excluded projects had lower scores, and were mostly older projects that were funded when success rates were much higher. Finally, Lindner & Nakamura (2015), both from NIH's Center for Scientific Review, provide quantitative and qualitative analyses and conclude that "retrospective analyses of the correlation between percentile scores from peer review and bibliometric indices of the publications resulting from funded grant applications are not valid tests of the predictive validity of peer review at the NIH."

Chance may also play an important role in the selection of which proposals to fund, where chance is related to reviewer selection, timing, and the competitive context within which a particular proposal is considered (Cole, Cole, & Simon, 1981). More recent studies have also quantified this variability in grant peer review. Johnson (2008) showed how modeling of differences in NIH reviewer scoring patterns along with transfer of information from panels to the final ratings group could change set of funded applications by up to 25%. Graves et al. (2011) used a bootstrapping technique to show that 59% of funded proposals submitted to the National Health and Medical Research Council of Australia might not have been funded when random variability was introduced into the actual review scores.

Only a few studies have attempted to identify features that correlate with proposal success. For example, Mutz et al. (2015) found that project duration, number of reviewers and gender all had an effect on proposal success. Viner et al. (2004) found that proposal success was correlated with greater experience or longevity. Van den Besselaar & Sandström (2017) look at the cognitive distance between 40 reviewers and 400 applicants over four panels. In three of the panels, applicants who were cognitively closer were more likely to get an award. In one panel the opposite was true. Overall, however, these findings do not stand up across multiple studies. Moreover, these studies have focused on applicants rather than projects and, as a result, none have correlated proposals with topic-level data that correspond to the subjects of the proposals.

The single study that is closest to what we will report in this paper is the recent one by Hörlesberger et al. (2013). They correlated the funding results from 198 ERC proposals (of which 41 were funded) with five indicators of their own design, some of which rely on document clusters (i.e., topics) in the AIT model of science that was created from the PASCAL database:

- Timeliness – age of proposal references
- Similarity – innovativeness of the topical cluster to which the proposal linked using title/abstract terms
- Risk – overlap between proposal references and references in papers authored by the applicant

- Pasteuresqueness – combination of patent links and fundamental or applied nature of journals published in by the applicant
- Interdisciplinarity – breadth of extracted keywords over disciplines

They found modest but statistically significant correlations between proposal success and two indicators, 'similarity to frontier research' and interdisciplinarity, while their risk indicator had the highest correlation, but was not statistically significant.

Current Gaps

The empirical literature, to date, has focused on whether funding agencies are fair in terms of the decisions they make. However, there is little prior literature (i.e., a gap) on proposal success and failure that takes the perspective of the individual or institution that generates these proposals. Different questions emerge when one takes this perspective. Individuals have to decide which topics they should focus on. Do they fare better if they are a publication leader in the topic? Does it make a difference if the topic has suddenly become more popular? In addition to individual researchers, policy makers in research institutions have to make decisions that affect the portfolio of proposals. Deans might be hiring elite faculty because of an (erroneous) belief in the Matthew effect. Research offices might be providing additional administrative support to elite faculty because of a belief in the Matthew effect. What if the prior literature is correct and there isn't any solid evidence of a Matthew effect? Policies might be based on perceived advantages that don't, in fact, have any impact on proposal success and failure. These questions have not been addressed in the prior literature.

There is a second gap is literature that needs attention. There is a lack of empirical studies that focus on the clarity of a research proposal. On the one hand, it is not uncommon to find courses on research proposal writing or businesses that will write (or re-write) research proposals. There are books written on 'how to write a successful research proposal'. Program officers and researchers who participate in peer review panels claim that proposals that are difficult to understand are rarely funded. While many believe that clarity matters, we could not find any peer reviewed empirical studies that measure 'writing clarity' and correlate this statistic with proposal success or failure. The need to write clearly is a commonly held belief with no empirical testing in the context of proposal success and failure.

There are quantitative indicators of writing quality, such as Flesch scores and the Gunning fog index that have been used to show increases in the readability of medical articles during the editorial review process (Biddle & Aker, 1996; Roberts, Fletcher, & Fletcher, 1994). But of greater importance, in our opinion, is the clarity of the argument within the research proposal rather than the number of words per sentence. Thus, we build upon the work of John Swales[2], especially as his body of work emphasizes discourse analysis of scientific articles.

Discourse analysis is a broad term, and refers to the way that language is used. Swales, and those whose work builds upon Swales, such as Simone Teufel (2010), have applied these ideas to

---

[2] Professor Swales' two most highly cited works are entitled 'Genre analysis: English in academic and research settings (cited nearly 12,000 times in Google Scholar) and 'Research genres: Explorations and applications' (cited over 2,000 times). As an emeritus professor at Linguistics at the University of Michigan, he remains very active in the field.

research articles. We seek to apply discourse analysis to research proposals, where the aim of the document may be very different. For example, the central claim in most research articles is that the authors' discoveries are a contribution to knowledge – the metaphor of 'filling a gap in the literature' is a common means to communicate this claim. Research articles look backward in time to show what has been accomplished. In contrast, research proposals are stating an intention about the future – that the applicant can solve an important problem if funding is provided.

Data and Research Design

Given that the prior literature suggests that eminent scientists don't have an inherent advantage in terms of proposal success, funding decisions might, in fact, be random. One might therefore wonder if looking for ways to predict proposal success is doomed from the start. Our answer to this is no. In most situations where funding is competitively awarded there will be a set of proposals that are simply not competitive for some reason. If these reasons can be linked to characteristics of the proposals or applicants in some fashion, even partially, then indicators can be developed that might predict proposal success.

We analyzed 369 full-text, R01 proposals submitted to NIH in calendar year 2010 by UMMS faculty. R01 grants are typically used to support discrete, specific research projects. Funding rates for these grants are shown in Table 1, along with the corresponding funding rates for all R01 proposals submitted to NIH that year. A large majority (290) of the 369 proposals were new, while the remaining (79) proposals were applications for renewal of an existing grant. The overall funding rate for this sample of proposals (22.8%) was only slightly higher than the NIH norm (22.0%).

Table 1. Proposal types and funding success rates in 2010 for UMMS R01 proposals.

Proposal Type	Total UMMS Submitted	UMMS Funded (Success Rate)	NIH R01 (2010) Success Rate
New	290	60 (20.7%)	17.8%
Renewal	79	24 (30.4%)	38.0%
Overall	369	84 (22.8%)	22.0%

Our research design is represented in Figure 1. First, the PI names and reference lists were extracted from the PDF proposal documents. We then linked each reference with a Scopus article ID (where possible) using reference matching techniques, and identified the topic associated with each reference using the document-to-topic assignments from the SciTech Strategies (STS) model of science, which consists of over 58 million documents assigned to 91,726 topics (Klavans & Boyack, 2017). This allowed us to fractionally assign each proposal to a set of topics using the proposal references.

The average proposal in our sample had 80.8 references (see Table 2), which is roughly double the number of references in a typical biomedical research paper. Renewals had, on average, a few more references than new proposals. We found that the large majority of references in all proposals could be identified as either a source paper or non-source paper in our model of

science, and were thus able to identify the topic associated with each of the linked references. This allowed us to identify the distribution of topics that were referred to by each proposal and, ultimately, allowed us to fractionally assign each proposal to topics. We calculated a Herfindahl (or concentration) index for each proposal from the distribution of references over topics, where the Herfindahl is the sum of the squared fractions. Proposals range from referencing a very large and dispersed set of topics (Herf ≈ 0) to having nearly all references in a single topic (Herf ≈ 1).
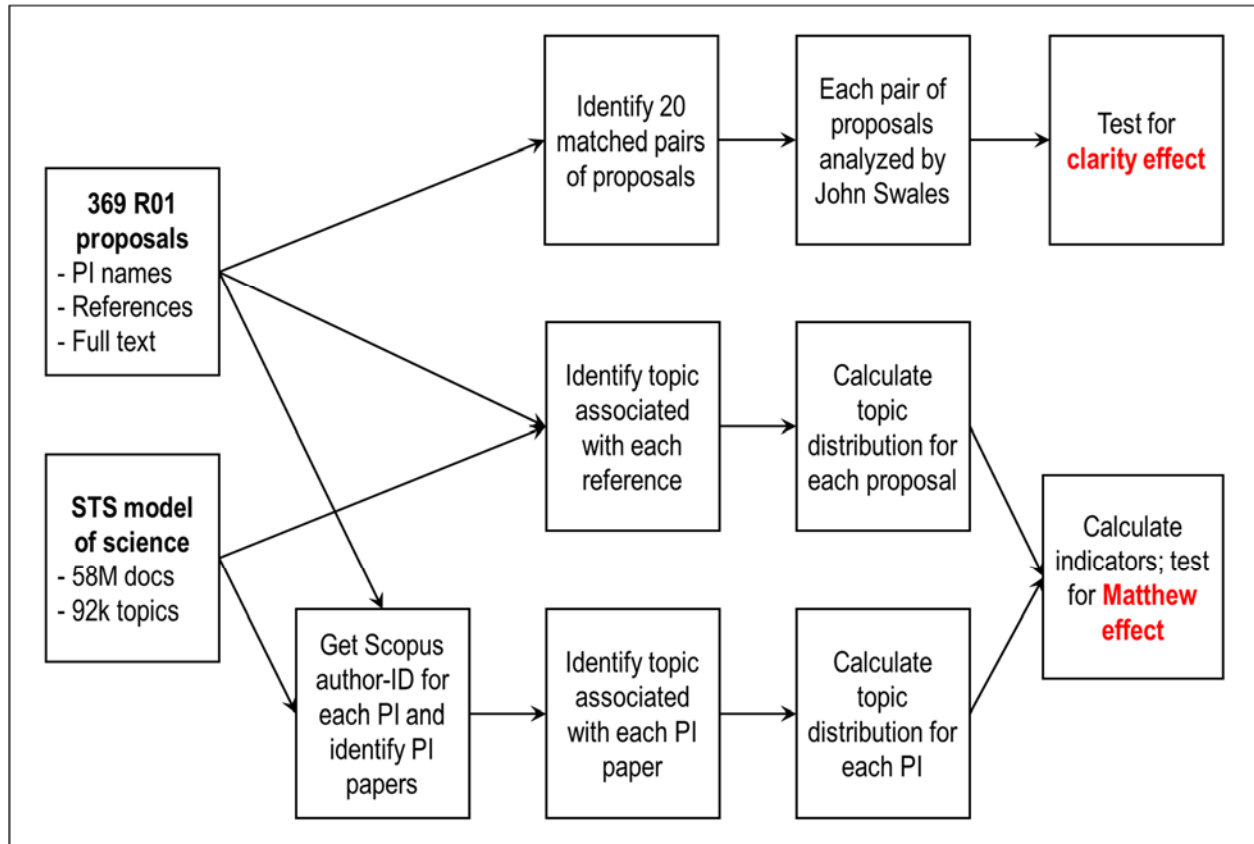

Figure 1: Overview of the research design

Table 2. Proposal reference and applicant publication ranges (means in parentheses).

	New proposals	Renewal proposals
# Proposal references	2 – 282 (80.8)	19 – 199 (86.7)
Herfindahl (refs)	0.022 – 0.946 (0.151)	0.032 – 0.505 (0.175)
# Applicant publications	2 – 487 (91.7); 223 PI	20 – 380 (123.0); 70 PI
Herfindahl (pubs)	0.023 – 0.699 (0.166)	0.023 – 0.640 (0.198)

We also identified Scopus author IDs for all but one of the 274 unique principal investigators (PI), and then identified the papers that had been written by each PI prior to 2010. Using the topic assignments for these papers, we then fractionally assigned each PI to a set of topics. Of the applicants, 199 submitted one proposal, 59 submitted two proposals, and 16 submitted either

three or four proposals. Applicants of new proposals had an average of 91.7 published papers as of 2010. Thus, the majority of proposals were submitted by experienced researchers.

We also used the STS model of science to calculate the relative publication share (RPS) of the UMMS in each topic. For those topics where UMMS has more papers than any other U.S. institution, RPS is defined as the number of papers by the University of Michigan (the entire university and not just the medical school) divided by the number of papers by its closest competitor, and is greater than 1.0. For those topics where another institution has more papers, RPS is the number of papers by the University of Michigan divided by the number of papers by the top institution in that topics, and is thus between 0.0 and 1.0.

Our first three hypotheses focused on the Matthew Effect.

H1: We tested whether proposals from applicants with higher productivity will be more successful than proposals from applicants with lower productivity. The total number of PI publications (up until the date of the proposal submission) was used as the indicator of PI productivity.

H2: The second test looks at the match between the topic profile of the proposal and the topic profile of the PI. We are, in essence, determining if the applicant has strengths in the topic associated with the proposal. Figure 2 shows two examples of overlap between proposal references and applicant publications – one with relatively low overlap, and one with a much higher overlap. In the low overlap case, references and publications are spread over 100 topics. Roughly 30% of the references and publications are in those few topics (in the middle of the chart) that contain both references and publications, while only 12% of references and 7% of publications are in the dominant topic. For the high overlap case, over 50% of references and publications are in the overlap space, while 38% of references and 32% of publications are in the dominant topic.
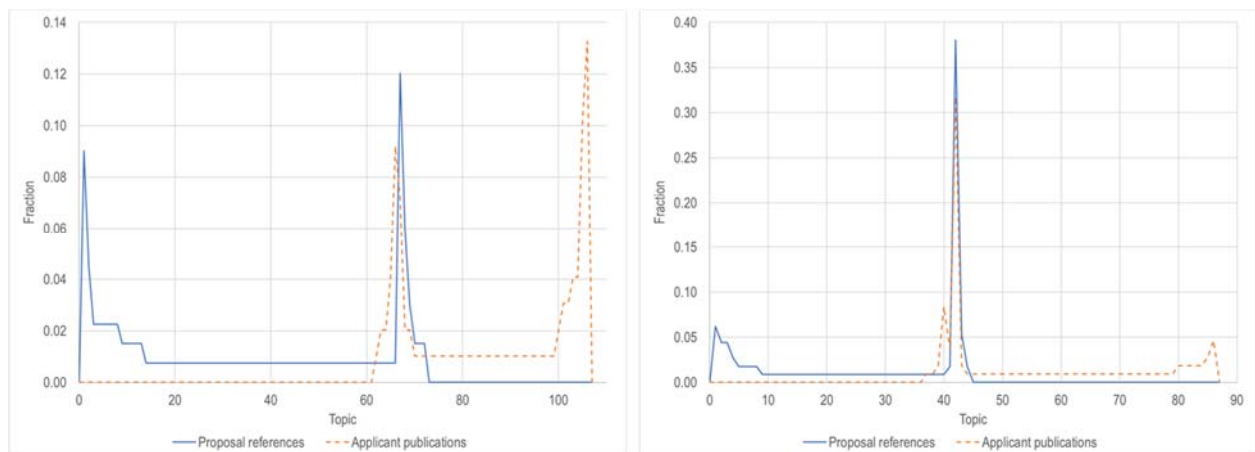

Figure 2. Representation of cases of low overlap (left) and high overlap (right) between proposal references and applicant publications for two of the proposals in our sample.

The indicator for H2 is calculated as Alignment = $\sum$ Sqrt ($P_{i,ref} * P_{i,pub}$), summing over all topics, where $P_{i,ref}$ and $P_{i,pub}$ are the percentages of proposal references and applicant publications by topic i, respectively. For the two cases in Figure 2, the values of this indicator are 0.272 and 0.476, respectively.

H3: This hypothesis is based on the assumption that a university is more likely to receive funding in topics where it is strong rather than weak. Here we define strength as having a high (within topic) publication share relative to other U.S. universities. The indicator for H3 is calculated as Strength = $\sum$ ($P_{i,ref} * RPS_i$) where $P_{i,ref}$ is as defined above (and shown in Figure 2), and RPS_i is the relative publication share for the institution in topic i. To illustrate this hypothesis, if a particular proposal was focused on two topics, and the university had the largest number of publications in those two topics of any U.S. institution, the RPS values would be very high, and one might expect the proposal to be funded because of historical strengths in these topics.

Figure 1 also shows a pathway to test for the clarity of proposals. For this test, we identified 20 pairs of proposals where the (within pair) topics were similar and the (within pair) PIs had similar H-indexes. Bibliometric data were used to inform the choosing of these pairs. Only one of the proposals in each pair was successful. We then asked John Swales, Professor Emeritus of Linguistics at the University of Michigan, to read each pair of proposals and tell us which proposal in each pair had a higher quality of discourse. Prof. Swales was blinded to investigator name, gender, ethnicity, and proposal success. He was instructed to ignore content and make his determinations based solely on the quality of discourse.

H4: Hypothesis four is based on Prof. Swales' judgment about the relative quality of the discourse for each of the 20 matched pairs of proposals. If he judged both proposals in a pair to have writing of equal quality then both proposals were given the value of 0.5. If he judged one proposal in a pair to have a higher quality of discourse than the other proposal, the higher quality proposal was given a value of 1.0 and the lower quality proposal was given a value of 0.

Results

The results of our analysis are shown in Table 3. For purposes of comparison, we have converted all statistical tests into simple Pearson correlation statistics. The dependent variables are success/failure (1,0) in the submittal period and success/failure (1,0) in any year. We chose not to analyze new proposals and renewals together because they are reviewed differently, and have differing success rates. The results here are for new proposals only; there were too few renewals to achieve statistical significance on any of our tests. A detailed discussion of results related to each hypothesis is given below.

Table 3. Summary of statistical findings

	Feature	Correlation	#Observations
H1	# Applicant publications	-0.013	290
H2	Applicant-proposal alignment	0.163**	290
H3	Institution research strength	-0.012	290
H4	Quality of the argument	0.474**	40

significance @ *0.10, **0.01

H1: Applicant productivity is not a predictor of proposal success. In this relatively homogeneous sample of 290 proposals, the prior publication record of the applicant didn't matter. This correlates well with previous studies (Bornmann, et al., 2010; Hornbostel, et al., 2009; Melin & Danell, 2006; Neufeld & Hornbostel, 2012; Neufeld, et al., 2013; Saygitov, 2014; Van den Besselaar & Leydesdorff, 2009) that have found no association between proposal success and applicant productivity.

H2: Proposals are more likely to be successful when there is a match between the topic profiles for the proposal references and the applicant's previous publications. Stated more simply, proposals are more likely to be funded when there is evidence that the applicant has strengths in the topic area associated with the proposal. This is not altogether surprising in that NIH is often thought to fund safe, rather than innovative, research (Nicholson & Ioannidis, 2012). This finding is also similar to the results of Hörlesberger et al. (2013), who found a relatively high correlation between proposal success and the overlap between their references and the references in PI publications.

H3: Institutional research strengths do not predict proposal success rates. This finding is controversial because the common wisdom is that there is value in building up research strengths in a university. The data does not support this commonly held belief. According to these data, the reputation of a university for leadership in a specific topic (based on their publications) doesn't improve proposal success rates.

H4: Professor Swales considered 12 of the matched pairs to be of equal quality. Of the remaining eight pairs of proposals, he correctly predicted seven out of the remaining eight pairs based on their having a higher quality of discourse. Although this is a relatively small sample (only 40 research proposals), the strength of the correlation suggests that discourse analysis has the potential to be a good predictor of proposal success in larger studies.

Discussion and Implications

Rhetoric may be more important than references. Work on argumentative zoning (Teufel, Siddharthan, & Batchelor, 2009), which has been applied to the full text of scientific documents, can identify how a successful argument is constructed. There are a host of related techniques and best practices associated with the development and analysis of stories that might be applied to the appropriate sections of research proposals. A plethora of prescriptive literature exists on this topic and is widely available. From Strunk and White's classic *The Elements of Style*, to more recent contributions like Fish's *How to Write a Sentence: And How to Read One*, or even Steven King's *On Writing* – the explosive growth of M.A. and M.F.A. programs in Creative Writing over the last 30 years has ensured that there are a great number of writing manuals in circulation authored by widely respected masters of the craft.

While citation analysis has yielded far fewer results, the finding that proposal success is correlated with the overlap between proposal references and PI publications also goes to the heart of the problem of evaluating funding agencies. On the surface, one might want an agency to fund the best proposals on important topics, including those that may appear to be risky.

However, one could also ask other questions. For instance: Who defines which problems are important? How much funding should go to elegantly argued proposals in topics that may seem less important to the funder? How much funding should go to poorly articulated proposals on what are considered to be very important topics? In our opinion, funding research on the right problems is perhaps more important than funding research on established, but perhaps less important, topics with elegantly framed proposals by researchers with decades of experience.

All of our conclusions are tentative because of the small sample size of this proof-of-principle study. Furthermore, while some of the results are statistically significant in this small sample, the results only apply to a single applicant (UMMS), a single type of grant mechanism (R01), and a single agency (NIH). There is no guarantee that these results can be replicated or will be supported with a larger sample. However, the results are striking enough to suggest that scale-up of these activities is justified.

In addition, we think it is reasonable to draw out the following three implications of expanding proposal analytics. First, it is fair to claim that the benefits of creating and analyzing a large database of research proposals are far greater than the costs. The potential benefits are both tangible and intangible. Tangible benefits include the possibility for an institution or researcher to increase success rates and funding. Intangible benefits start with the possibility of analyzing the intentions (and hoped-for research strengths) associated with proposals in addition to the retrospective analysis of strengths from publications. Our experience at UMMS suggests that the costs associated with proposal analysis are relatively low. A university already owns the intellectual rights to their proposals (Funding agencies cannot analyze these proposals for reasons other than deciding on whether to fund the proposal without violating the intellectual rights of the university owners). The costs of converting pdf files into a structured database are inconsequential relative to the possible gains in grant funding that can be achieved by analyzing these proposal data.

Second, we claim that insights into research intentions will allow the field of scientometric analysis to expand its scope of application. The current scope is aimed almost entirely at research evaluation using data on scientific publications. This is appropriate since publications are strategic outputs that can be measured. However, as pointed out by Mintzberg & Waters (1985), there is a significant difference between strategic intent and strategic outcomes. The analysis of research proposals will allow the field of scientometric analysis to contribute to our understanding and articulation of strategic intent.

Finally, in order to exploit this scientific opportunity, the field of scientometric analysis may need to shift its focus from references to rhetoric. References are used to support rhetorical claims in scientific documents, but they don't seem to play the same role in research proposals. They may, for instance, simply be a means of attempting to convince a review panel that the applicant knows the topic well. Techniques of rhetorical analysis will need to be developed and tested. We don't question that this is possible. We posit that it will simply take time, effort, and the creative application of textual analytic tools.

Our understanding of the 'science of science' has significantly benefited from electronic access to the scientific literature. Specifically, the creation of a citation database by Eugene Garfield

(1955) and the corresponding development of different forms and methods for citation analysis have created a better understanding of the evolution of research supply. However, our knowledge of the evolution of the demand for science, as poignantly pointed out 10 years ago by Sarewitz & Pielke (2007), is marginal at best. Their proposed 'big data' method for characterizing the demand for science failed[3]. We need methods to identify and measure the various types of demands that are driving the science system, and we suggest that analyses of proposal data, perhaps using discourse analysis, has the potential to fill this need, just as Garfield's citation indexes met the needs he identified so many years ago.

References

Biddle, C., & Aker, J. (1996). How does the peer review process influence AANA Journal article readability? *Journal of the American Association of Nurse Anesthetists, 64*(1), 65-68.

Bornmann, L., & Daniel, H.-D. (2005). Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees decisions. *Scientometrics, 63*(2), 297-320.

Bornmann, L., & Daniel, H.-D. (2006). Selecting scientific excellence through committee peer review – A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics, 68*(3), 427-440.

Bornmann, L., Leydesdorff, L., & van den Besselaar, P. (2010). A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. *Journal of Informetrics, 4*, 211-220.

Bornmann, L., Wallon, G., & Ledin, A. (2008). Does the committee peer review select the best applicants for funding? An investigation of the selection process for two European Molecular Biology Organization programmes. *PLoS ONE, 3*(10), e3480.

Cabezas-Clavijo, A., Robinson-Garcia, N., Escabias, M., & Jimenez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *PLoS ONE, 8*(6), e68258.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science, 214*, 881-886.

Cole, S., Rubin, L., & Cole, J. R. (1978). Peer review in the National Science Foundation: Phase One of a study. In.

Enger, S. G., & Castellacci, S. (2016). Who get Horizon 2020 research grants? Propensity to apply and probability to succeed in a two-step analysis. *Scientometrics, 109*, 1611-1638.

Fang, F. C., Bowen, A., & Casadevall, A. (2016). NIH peer review percentile scores are poorly predictive of grant productivity. *eLife, 5*, e13323.

Gallo, S. G., Carpenter, A. S., Irwin, D., McPartland, C. D., Travis, J., Reynders, S., Thompson, L. A., & Glisson, S. R. (2014). The validation of peer review through research impact measures and the implications for funding strategies. *PLoS ONE, 9*(9), e106474.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science, 122*, 108-111.

Graves, N., Barnett, A. G., & Clarke, P. (2011). Funding grant proposals for scientific research: Retrospective analysis of scores by members of grant review panel. *British Medical Journal, 343*, d4797.

---

[3] Personal communication with Dr. Sarewitz on October 6th, 2017.

Herbert, D. L., Barnett, A. G., Clarke, P., & Graves, N. (2013). On the time spent preparing grant proposals: An observational study of Australian researchers. *BMJ Open, 3*, e002800.

Hörlesberger, M., Roche, I., Besagni, D., Scherngell, T., Francois, C., Cuxac, P., Schiebel, E., Zitt, M., & Holste, D. (2013). A concept for inferring 'frontier research' in grant proposals. *Scientometrics, 97*, 129-148.

Hornbostel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & Von Ins, M. (2009). Funding of young scientist and scientific excellence. *Scientometrics, 79*(1), 171-190.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics, 95*(9), 1168-1177.

Johnson, V. E. (2008). Statistical analysis of the National Institutes of Health peer review system. *Proceedings of the National Academy of Sciences of the USA, 105*, 11076-11080.

Klavans, R., & Boyack, K. W. (2017). Research portfolio analysis and topic prominence. *Journal of Informetrics, in press*.

Li, D., & Agha, L. (2015). Big names or big ideas: Do peer-review panels select the best science proposals? *Science, 348*(434-438).

Lindner, M. D., & Nakamura, R. K. (2015). Examining the predictive validity of NIH peer review scores. *PLoS ONE, 10*(6), e126938.

Melin, G., & Danell, R. (2006). The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy, 33*(10), 702-712.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal, 6*, 257-272.

Mutz, R., Bornmann, L., & Daniel, H.-D. (2015). Testing for the fairness and predictive validity of funding decisions: A multilevel multiple imputation for missing data approach using ex-ante and ex-post evaluation data from the Austrian Science Fund. *Journal of the Association for Information Science and Technology, 66*(11), 2321-2339.

Neufeld, J., & Hornbostel, S. (2012). Funding programmes for young scientists - Do the 'best' apply? *Research Evaluation, 21*, 270-279.

Neufeld, J., Huber, N., & Wegner, A. (2013). Peer review-based selection decisions in individual research funding, applicants' publication strategies and performance: The case of ERC Starting Grants. *Research Evaluation, 22*, 237-247.

Nicholson, J. M., & Ioannidis, J. P. A. (2012). Conform and be funded. *Nature, 492*(7427), 34-36.

Reinhart, M. (2009). Peer review of grant applications in biology and medicine: Reliability, fairness and validity. *Scientometrics, 81*(3), 789-809.

Roberts, J. C., Fletcher, R. H., & Fletcher, S. W. (1994). Effects of peer review and editing on the readability of articles published in Annals of Internal Medicine. *Journal of the American Medical Association, 272*(2), 119-121.

Sarewitz, D., & Pielke Jr., R. A. (2007). The neglected heart of science policy: Reconciling supply of and demand for science. *Environmental Science & Policy, 10*, 5-16.

Saygitov, R. T. (2014). The impact of funding through the RF President's Grants for Young Scientists (the field - Medicine) on research productivity: A quasi-experimental study and a brief systematic review. *PLoS ONE, 9*(1), e86969.

Swales, J. (1986). Citation analysis and discourse analysis. *Applied Linguistics, 7*(1), 39-56.

Teufel, S. (2010). *The structure of scientific articles: Applications to citation indexing and summarization*. Stanford, CA: CSLI Publications.

Teufel, S., Siddharthan, A., & Batchelor, C. (2009). Towards discipline-independent argumentative zoning: Evidence from chemistry and computational linguistics. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 1493-1502). Singapore.

Van den Besselaar, P., & Leydesdorff, L. (2009). Past performance, peer review and project selection: A case study in the social and behavioral sciences. *Research Evaluation, 18*(4), 273-288.

Van den Besselaar, P., & Sandström, U. (2015). Early career grants, performance, and careers: A study on predictive validity of grant decisions. *Journal of Informetrics, 9*, 826-838.

Van den Besselaar, P., & Sandström, U. (2017). Influence of cognitive distance on grant decisions. In *Science, Technology and Innovation Indicators 2017*. Paris, France.

Van Leeuwen, T. N., & Moed, H. (2012). Funding decisions, peer review, and scientific excellence in physical sciences, chemistry, and geosciences. *Research Evaluation, 21*, 189-198.

Viner, N., Powell, P., & Green, R. (2004). Institutionalized biases in the award of research grants: a preliminary analysis revisiting the principle of accumulative advantage. *Research Policy, 33*(3), 443-454.

Von Hippel, T., & Von Hippel, C. (2015). To apply or not to apply: a survey analysis of grant writing costs and benefits. *PLoS ONE, 10*(3), e0118494.

Zuckerman, H. (1967). Nobel laureates in science: Patterns of productivity, collaboration, and authorship. *American Sociological Review, 32*(3), 391-403.