Research Evaluation **24** (2015) pp. 271–281
Advance Access published on 4 May 2015
doi:10.1093/reseval/rvv009

Defining the role of cognitive distance in the peer review process with an explorative study of a grant scheme in infection biology

Qi Wang[1,*] and Ulf Sandström[1,2]

[1]*INDEK, KTH-Royal Institute of Technology, Lindstedtsvägen 30 114 28 Stockholm, Sweden and*
[2]*School of Business, Örebro University Swedish Business School, 701 82 Örebro, Sweden*
**Corresponding author. Email: qiwang@kth.se*

The aim of this article is twofold: (1) to provide a methodology for measurement of cognitive distance between researchers and (2) to explore the role of cognitive distance on the results of peer review processes. Cited references and the content of articles are used to represent their respective scientific knowledge bases. Based on the two different approaches—Author-Bibliographic Coupling analysis and Author-Topic analysis—we apply the methodology on a recent competition for grants from the Swedish Strategic Foundation. Results indicate that cognitive distances between applicants and reviewers might influence peer review results, but that the impact is to some extent at the unexpected end. The main contribution of this article is the elaboration on the relevance of the concept of cognitive distance to the issue of research evaluation in general, and especially in relation to peer review as a model used in grant decisions.

Keywords: cognitive distance; peer review process; grant decisions; author bibliographic coupling; author topic analysis.

1. Introduction

Peer review is intended to improve both the technical quality of projects in research and the credibility of the decision-making process. Nowadays it is taken for granted that peer review is fundamental to the institution of science and a symbol for the autonomy of science (Chubin and Hackett 1990). Although peer review functions are put into action to enhance the quality of research and to prevent poor research from taking place, the procedures do not always function as expected. Bias in peer review is a crucial issue that has generated serious discussions over a period of years (Wesseley 1998; Bornmann and Daniel 2005; Bornmann 2011). Any type of bias would be detrimental to the pursuit of excellent research at different research fronts.

Many possible flaws in the peer review process have been disclosed over the recent years. McCullough (1989) reported in a survey of principal investigators applying to the US National Science Foundation (NSF) during 1985, based on 9,500 respondents, that two-fifths were unsatisfied with the assessment of their proposals. Reasons for dissatisfaction were statements like: ‘reviewers or panelists are not expert in the field, poorly chosen, or poorly qualified’ (McCullough 1989). In a peer review process, reviewers are supposed to be experts in the field; however, the expertise and authority of reviewers are frequently being questioned. In fact, the guidelines for selecting reviewers in many journal editorial offices or research grant agencies are ambiguous. Occasionally project managers or editors select reviewers based on their experiences or personal relations (Caelleigh, Shea, and Penn 2001). If a reviewer is not an expert in the area under evaluation, the decision given might be unreliable or open for discussion.

Cognitive bias, also known as ‘cognitive particularism’ and ‘cognitive similarity’ (Travis and Collins 1991), refers to a situation where scientists with a mainstream view of their respective fields could pose challenges to a fair review process of new and alternative research strategies. Moreover, cognitive bias is generated because of the existence of cognitive boundaries within and between scientific specialties and disciplines (Travis and Collins 1991; Whitley 2000). Due to the difficulties of measuring the cognitive distance between applicants and reviewers, cognitive bias in the peer review process is often ignored. However, this bias might have a substantial effect on interdisciplinary research proposals because that type of research is often located at the boundaries of traditional disciplines, causing difficulties in finding suitable reviewers.

To fill this gap, this article discusses the role of cognitive distance in a peer review process by proposing an advanced measurement of cognitive distance between individual applicants and their reviewers and by evaluating to what extent cognitive distance impacts on peer review. The structure of this article is as follows. First, we provide background information and review of the concept of cognitive distance. Following that, we conceptualize the cognitive distance between applicants and reviewers. Next, we discuss and present a methodology based on Author-Bibliographic Coupling analysis (ABC) and Author-Topic analysis (A-T). In the following section, we report the results. We conclude by discussing the results and the advantages and shortcomings of the proposed methodology.

2. Research background

Few studies have investigated cognitive bias in peer review. A pivotal contribution by Mahoney (1977) found that 'reviewers were strongly biased against manuscripts which reported results contrary to their theoretical perspective'. In other words, it implies that reviewers would likely support manuscripts similar to their own. Later, Travis and Collins (1991) coined the terms 'cognitive particularism' or 'cognitive cronyism/similarity' to describe the different peer review situations. They believed that cognitive bias is caused by the 'cognitive structure of science' and that it 'depends on the existence of cognitive boundaries within and between scientific specialties and disciplines' (Travis and Collins 1991). Moreover, they made direct observations within a grant-awarding committee of the British Science and Engineering Research Council. With this qualitative method, they were able to indicate the effects of cognitive cronyism/similarity on peer review results. Meanwhile, they operationalized 'cognitive similarity' into measures for the department status of applicants and reviewers and their social positions. However, no clear conclusions were drawn from their fieldwork because the authors neither mentioned how widespread cognitive cronyism is nor specified how damaging it might be to the peer review.

Based on Travis and Collins, Sandström (2009b) developed a strategy for empirical investigation of cognitive bias. He introduced the concept of 'cognitive distance' in the peer review process, and proposed bibliographic coupling as a method to detect cognitive bias. The method was applied to a large grant scheme of the Swedish Research Council: the Linneaus Grants initiated in 2005. Preliminary conclusions based on mapping of applicants and their relations to reviewers indicated that groups who were not rewarded had fewer connections to reviewers than the granted groups. At the same time, it could be shown that the non-rewarded groups exhibited better results in track records using relative citation scores. Another research by Sandström and colleagues (2010) indicated that 'it was decisive to have a cognitive similarity in order to receive an excellent grant'. Out of three large calls for excellence grants, all groups that were granted had higher similarity compared to those not granted.

Full and detailed data on grant peer review are seldom disclosed due to secrecy and other policy issues. Two studies based on detailed data including bibliometric analysis have been published and both are Swedish: Wennerås and Wold (1997) and a follow-up 10 years later by Sandström and Hällsten (2008). These studies were made possible due to the Swedish principle of public access to official documents. But, beside gender and conflict of interest, these studies did not investigate the issue of cognitive distance, although other possible biases were covered.

In short, former strategies for measuring cognitive bias have been based on the information concerning applicants and reviewers such as departments, co-authorships, and cited references; surprisingly there are no studies focusing on the research content itself. In this study, we use the term 'cognitive bias' to interpret the bias caused by heterogeneity of theoretical perspectives among individual researchers and to explore the role of cognitive bias in peer review. In doing so, a strategy that combines measuring research tradition and content is applied to obtain the cognitive distance between applicants and referees.

Before entering into the conceptualization of cognitive distance, it is necessary to distinguish between manuscript peer review and grant peer review. In the former case, it should be easier to find relevant reviewers—for example, based on the manuscript references—but in the latter case, this is not possible because panels in standing committees have to be organized over a longer period of time. This makes the process sensitive to differences between research trails in fields which may have several traditions. If there are many different trails, there cannot be representatives for all because the committee membership is limited to six or seven members. Consequently, there is much more room for cognitive bias in the panel- or committee-organized peer review. It should be said that it is possible to combine several review approaches, as is the case in the NSF and in many other national research councils. The Swedish Research Council has worked for decades on the basis of nationally organized committees, but lately there seems to be a change toward more of international and mail peer review in combination with standing committees.

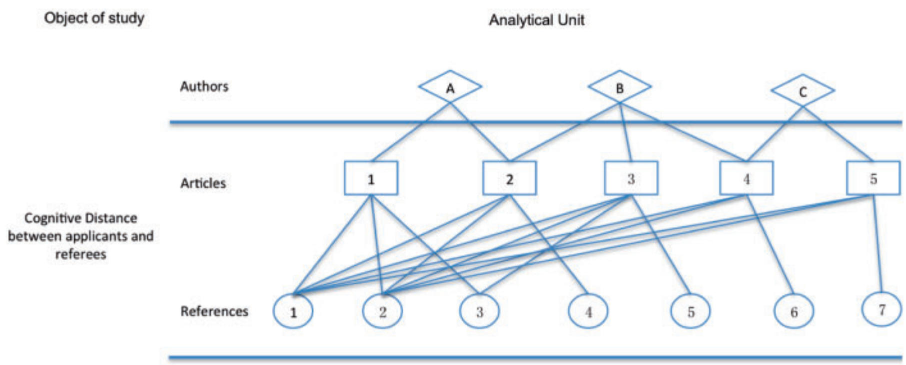

Figure 1. Scheme of operationalization of cognitive distance.

3. Conceptualization of applicant-reviewer cognitive distance

A vague concept could lead to misunderstandings, thus it is necessary to define precisely what we consider as cognitive distance between individual researchers. Cognition refers to 'a series of mental activities, including proprioception, perception, sense making, categorization, inference, value judgments, emotions, and feelings, which all build on each other' (Nooteboom et al. 2007). Hence, to measure the individual cognition seems an almost unmanageable task. However, when individuals are labeled as researchers, the cognitions laid bare by their scientific work are our only concern. Nooteboom and colleagues (2007) have stated that cognitive differences between individuals are the result of their respective knowledge bases. Here, we aim to take his conceptual work a bit further. In several papers by Nooteboom and others (Nooteboom 1999, 2000; Nooteboom et al. 2007), cognitive differences at the company level have been analyzed by utilizing patent data as a proxy for companies' knowledge base—that is, when two companies have (one or more) patents in the same category, it indicates a smaller cognitive distance between the two companies (Wuyts et al. 2005; Cantner, Meder, and ter Wal 2010; Dangelico, Garavelli, and Petruzzelli 2010).

For a researcher, the knowledge base might be the result of diverse sources, such as educational background, books or articles read, and research programs implemented. Here, considering data availability and quality, we use a researcher's cited references as an indicator of his/her research trail. The reason is that the cited papers are used to develop a researcher's own articles, and we assume that the research work of an individual researcher is highly related to the cited papers. Thus, we infer that the more references shared by the different authors, the smaller the cognitive distance between the two researchers.

Additionally, the research trail, or the research content itself, should reflect a researcher's cognition more directly. Researchers demonstrate how they understand, analyze, and interpret different problems through their research outcomes (text) in publication channels as journal publications, proceedings papers, reports, books, and patents. Thus, we could obtain the cognitive distance between researchers by measuring and comparing their research contents. Considering the efficiency of calculation, we use text from titles and abstracts obtained from the Thomson Reuters database Web of Science (WoS), instead of full text of papers. Accordingly, we assume that aggregating titles and abstracts from all of a researcher's publications would approximately represent this researcher's cognition. Fig. 1 summarizes the relations between research trail and research content.

4. Operationalization of cognitive distance

The proposed method for measuring the cognitive distance between applicants and reviewers can be subdivided into two perspectives: the first is ABC, and the second is the A-T.

4.1 Author bibliographic coupling analysis

ABC (Sandström 2008; Zhao and Strotmann 2008a, b; Ma 2012), which is an extension of the concept of bibliographic coupling (Kessler 1963), can be used to measure the knowledge similarity between researchers, to construct the intellectual structure of research areas, and even to represent the knowledge absorption, diffusion, flow of the research area, and so forth (Glänzel and Czerwon 1996; Boyack, Klavans, and Börner 2005).

Taking individual researchers as the study target, we have grouped the publications and references of each researcher. The relations among author, publication, and cited references that are shown in Fig. 1, can be represented by the following Table 1. As mentioned above, we did not exclude the duplicated publications in our data set; thus publications 2 and 4 appear more than once in Table 1.

Table 1. Description of bibliometric information shown in Fig. 1

Author	Publication	Reference
A	Publication 1	Ref 1; Ref 2; Ref 3
	Publication 2	Ref 1; Ref 2; Ref 4
B	Publication 2	Ref 1; Ref 2; Ref 4
	Publication 3	Ref 1; Ref 2; Ref 3; Ref 5
	Publication 4	Ref 1; Ref 2; Ref 6

Table 2. The Author–Reference matrix

Reference	Author A	Author B
Reference 1	2	3
Reference 2	2	3
Reference 3	1	1
Reference 4	1	1
Reference 5	0	1
Reference 6	0	1

Furthermore, we could create an author–reference matrix for ABC, which is shown in Table 2. It displays the cited times of each reference by individual researchers. Taking authors A and B as an example, they have published 2 and 3 documents, respectively, and one of the documents is their cooperative work. We added the references cited by the collaborative paper by both authors.

With the author-reference matrix as an input, the Salton's cosine (Salton and McGill 1983) was used to measure the similarity between applicant a and referee b. The formula is as follows,

$$cos(a, b) = \frac{\sum c_{ai}c_{bi}}{\sqrt{\sum c_{ai}^2}\sqrt{\sum c_{bi}^2}}.$$

Using this function, we obtain the similarity that is in the interval between 0 and 1. Then, the cognitive distance based on author bibliographic coupling can be calculated by

$$cog\_distance_{biblio} = 1 - cos(a, b)_{biblio}.$$

Obviously the smaller the cognitive distance is, the more similar their research would be. Furthermore, when the distance between an applicant and a referee is 1, it indicates that they have 0 references in common in their previous research. In other words, they differ in their research traditions, trails, or paths. On the contrary, if the cognitive distance is 0, it implies that all of the references are the same on both sides (applicant and reviewer). That might be a result of intense collaboration and jointly published papers.

There are several reasons for applying the author-bibliographic coupling method to test the cognitive similarity instead of other similar approaches, such as direct citation analysis and co-citation analysis. First, there is a time lag in the co-citation analysis (Hopcroft et al. 2004; Shitaba et al. 2009), which implies the fact that a certain time interval is required for conducting co-citation analysis. In comparison, Author-Bibliographic Coupling is more sensitive to recent publications. Meanwhile, although direct citation could avoid the time effect, its accuracy in assessing the similarity is inferior to the bibliographic-coupling method (Ahlgren and Colliander 2009).

4.2 Author-topic model

To measure the cognitive distance between applicants and reviewers regarding their cognitive content, we apply an Author-Topic model (Rosen-Zvi et al. 2004), which is an extension of the *Latent Dirichlet Allocation* method (Blei et al. 2003), by including textual information into the model. It presents the multinomial distribution of each author over topics. The advantage of this model is that it uses 'a topic-based representation in order to model both the content of documents and the interests of authors' (Rosen-Zvi et al. 2004).

Here we used the text data from titles and abstracts of publications to represent research content, and furthermore applied the Author-Topic model to obtain the distribution of individual researchers over multiple topics. However, identifying an appropriate number of topics is one limitation inherent in this model. Generally, there are two ways to solve the problem: one is training parameters by minimizing the complexity of a sample data; another is to use the rule of thumb to approximately estimate the number of topics. In this case, we chose the latter and identified 20, 30, and 40 research topics respectively. We then calculated the similarity using the Salton's cosine (Salton and McGill 1983) based on each author-topic matrix—that is, the same way as above. Furthermore, the paired-sample T-test and the Pearson's correlation coefficient analysis were conducted to determine whether the difference between every two sets of similarity were significant. The results show that the difference is not significant but the correlation coefficient is high and significant for every two sets. Therefore, we emphasize that the number of topics does not have considerable impact on the following analysis. In this case, we used the similarity measured from 40 research topics. Finally, cognitive distance based on the A-T can be obtained by the following formula,

$$cog\_distance_{topic} = 1 - cos(a, b)_{topic}.$$

Likewise, cognitive distance obtained by the A-T is in the interval between 0 and 1. If an applicant and a referee have a small cognitive distance obtained by the model, it indicates that they are quite similar in the terms used in title and abstract. On the other hand, if the cognitive distance is

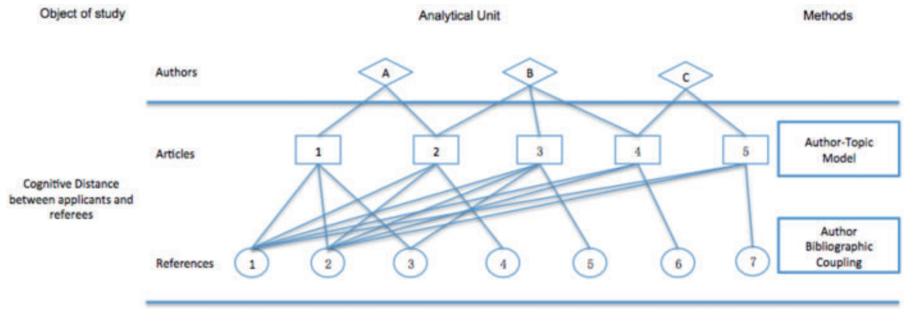

Figure 2. Summary of the proposed methodology.

large, it implies that they differ in their use of research terms.

4.3 Short summary on methodology

We propose a combined method to measure cognitive distance where both the references and the content of individual researcher and reviewer are considered. Previous research on this problem has paid little attention to the latter aspect, that is, research content; solely references were used (Sandström 2009b; Sugimoto and Cronin 2013). In our view, references could reflect research traditions/trails of an individual researcher. Furthermore, with the A-T, we could obtain cognitive distances in their research content itself. Studies in computer science have applied topic models to match submissions with referees (Mimno and McCallum 2007; Daud 2012). However, the drawback of this kind of technique is that it is difficult to detect the researcher's attitude on specific theoretical perspectives. Different schools probably differ in perspectives, interpretations, and research methods/paradigms to the same research question. For instance, in the case of classical economics, new classical economics, Keynesian economics, and the like, they all have the same focus in economics research but are extremely far from each other. Thus, it is quite important to have insight into a researcher's tradition/trail in order to be able to correctly classify the content of a paper. Because cognitive distances obtained by A-T and ABC, respectively, have different implications, we did not provide an integrated algorithm. Figure 2 summarizes the methods we proposed.

The strength of our method is that collaborative relations are adequately addressed. Obviously, collaboration is an important way to achieve cognitive similarity and absorptive capacity (Nooteboom 2000; Hautala 2011). The more collaborative work among researchers, the more similar their cognitive relations will be. In our measurement, if two researchers have active collaborations, the cognitive distance between them would be shorter than if they are only refereeing to the same references. If the distance is short without collaboration then we can infer that they are competitors at a specific research front.

5. Case and data

The case used in this article was initiated by the Swedish Foundation for Strategic Research (SSF). The full name of the scheme is 'Molecular mechanisms in the interplay between microorganisms/parasites and their host (man, domestic animals, plans and forest trees) in relation to disease'. In 2013, SSF decided to invest 225 million SEK on projects that would 'result in new knowledge that may be used in finding cures for malaria or cholera or in the development of new antibiotics, diagnostic tests or vaccines' (SSF 2013). Projects were organized as framework grants aiming at stimulating individual researchers, from both academic and industrial fields, to collaborate to conduct 'excellent' research.

The foundation received 57 research proposals, from 57 main applicants with 136 co-applicants. To select the projects with potential value, SSF used a two-stage peer review approach. In the first round, 14 referees involved had a diversity of backgrounds both from university and from the pharmaceutical industry. Referees were Swedish or Swedish permanent residents. Note that some of the referees in the first round had no publications; they were chosen from relevance criteria (industry representatives) with none or very few recent academic merits. This, of course, leads to difficulties in measuring cognitive distance.

Twenty-eight out of 57 applications advanced to the next stage. Unlike the previous round, nine international referees (non-Swedish) were selected (by whom the referees were selected was not disclosed by the foundation). Nine proposals were granted. A single-blind type of 'peer review' was applied in both rounds, which implies that referees could review the resumes of applicants, including the information on educational background, professional experiences, publications, and the like. It is highly probable that all the referees were involved in the review of each application. However, it is not clear whether referees could discuss or exchange views among themselves during the review process. It is unclear whether referees had an actual meeting in the same location.

Data on publications were collected from the WoS database SCI-E using the following document types: *Article*, *Letter*, *Proceeding Paper*, and *Review*. Names of applicants and referees were used to search and retrieve publications. This might have led to the obtainment of redundant publications due to duplicate names (homonyms). To make the data accurate, we refined data automatically based on all possible information, such as source, organization, and country. But due to collaborations among applicants and even between referees and applicants, there are a few duplicate publications in our data set. These duplicates were not removed. Finally, the total number of publications (not unique) obtained was around 8,000.

According to the regulations of SSF, every referee is likely to be involved in each application's review process. We thus measured cognitive distances between main applicants and each referee, and then used the minimum distance for each relation to represent the cognitive distance between an application and its possible referees. There are definitely other strategies that could have been adopted, for instance, measuring cognitive distances between each referee and every researcher within a group and then using the average cognitive distance. The reason of considering main applicants instead of every involved researcher is that, in this case, main applicants are probably the most prestigious persons in the group. In addition, we would like to emphasize that the choice should depend on the case per se. If the aim of a grant is to provide young researchers with an opportunity to conduct their research, young researchers then should be the focus of analysis.

6. Result

6.1 Result from first-round peer review

In the first review stage, there were 57 main applicants, of which two had no publications in WoS. Fourteen referees were involved in the review session. About half, 28 out of 57, of the applicants were forwarded to the next peer review round. Figure 3 shows the results. The horizontal axis represents the cognitive distance measured by A-T analysis, whereas the vertical axis shows results based on ABC analysis. Each dot represents an application, and the red color represents those applications that were forwarded to the second round, whereas the blue color is for the failed applications.

First, as shown from the vertical axis (measured by the ABC analysis) in the above figure, most applicants have large cognitive distances to reviewers, concentrated in the interval from 0.95 to 1. Long distances obtained by ABC analysis indicate that applicants and referees rarely use the same references.

Second, cognitive distances obtained by the A-T analysis were scattered from 0 to 1, and clearly there are only a few applicants with extremely short (0–0.2) or long (0.8–1) cognitive distances to their reviewers. Based on this result, we can infer that some of reviewers have few publications in their reviewed research areas. Therefore, it is difficult to ascertain whether or not the reviewers are experts in the field; actually, the credibility of the procedure for (*peer*) review in this case could be seriously doubted assuming that peers should be active in applicant's areas.

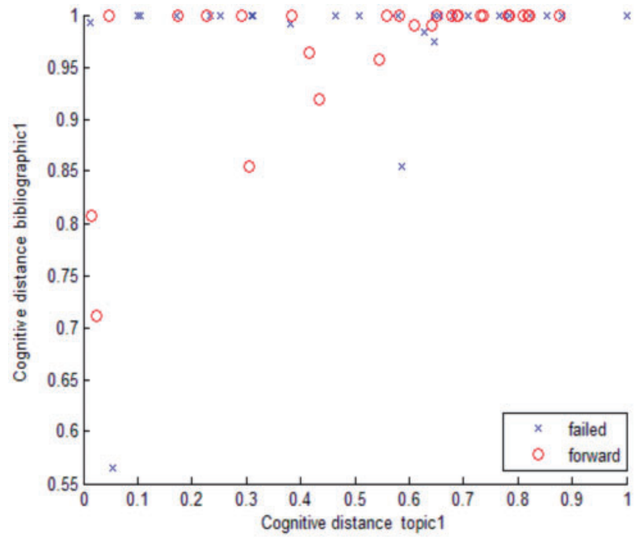

Figure 3. Cognitive distances between applicants and reviewers in the first round.

6.2 More tests on first-round peer review

As mentioned above, the reviewers involved in the first round were all from Sweden, and they were from many sectors of society. Thus, there might be a small country problem, which implies that 'personal relations and politics might dominate the scene and objective impartial evaluation is not possible' (Pouris 2007). In this case, other factors rather than the applications themselves could play an important role in the review process. Therefore, we investigate whether previous academic performance of applicants could have affected the review results. The individual academic performance is often evaluated by bibliometric indicators, such as the number of publications, journal impact factors, and academic positions. We have noticed that various bibliometric indicators are designed to describe research performance at the level of both individual researchers and research institutes, such as the CPP/FCS (citations per publication/field citation score) (van Raan 2005), MNCS (mean normalized citation score) (Waltman et al 2011), H-index (Hirsch 2005). In this study, we prefer to use simple and straightforward bibliometric indicators, since it is difficult for reviewers to conduct complex evaluation on the previous academic performance of applicants. The bibliographic indicators were selected mainly based on the considerations of quantity and quality of publications as well as the 'newness' of the research.

- The number of publications. It reflects the productivity of an individual researcher; the more publications one has, the more chance reviewers could have a deep impression on this researcher. Here we used the indicator of fractionalized papers (*Frac P*), which measures the relative number of publications of each applicant.

- The quality of publications. It is measured from two perspectives, the number of citations that publications have received and the citation score of journals where publications appeared. The reasons are twofolds. First, publications that appear in a journal with an excellent reputation imply that the publications should display good quality. The variable, *Top1%*, is used to measure the share of articles cited above the 99th percentile. Meanwhile, publications that do not appear in the journal with excellent reputation may also receive a great number of citations. To avoid differences caused by WoS subject categories, we used normalized journal citation score (*NJCS*) to assess the quality of journals rather than journal impact factor.
- The 'newness' of the research. The variable of 'vitality' is included, since we assume that the innovative character of applications might be an important indicator when reviewers decide whether applications should be supported.

We also included two control variables regarding the identities of applicants: gender, valuing 1 if the applicants are female, and scientific position, valuing 1 if the applicants are professors. The detailed measurement methods as well as descriptive statistics of each variable are provided in the appendix. Besides, more detailed explanations and measurement of the variables can be found in an evaluation report by Sandström (2009a).

We used a logistic regression model to determine what factors could be involved. Logit model is one of binary data models, which is used when the 'dependent variable can take only two possible values, say $y = 0$ or $y = 1$' (Cameron and Trivedi 2005). In our case, the dependent variable is whether applications entered to the second review round, which is binary and not continuous, thus a logit model is used. In this section, we should emphasize that data from all applicants including both principal investigators and co-applicants were used.

Table 3 reports the results of logistic regression, investigating the impact of bibliographic indicators on first round review results. First, *NJCS* and *Top1%* are significant and positive, indicating that the more articles published in good journals, the more possible it was that they could move forward to the next review round. Obviously, the quality of the journal has a strong impact on the review results. However, *Frac P* and *Vitality* are not significant. It seems that reviewers did not consider the number of publications of individual applicants. *Vitality* was used as an indicator of research novelty, which is not significant in the model; thus, it is not an influential factor in peer review (at least, not in this specific case under these specific circumstances).

Table 3. Logit regression in first round selection

Variable name	Coef.	S.E.
Frac P	0.0794	0.0575
NJCS	1.5988**	0.6458
Top1%	22.3602**	8.1110
Vitality	3.237	2.117
Gender	−0.2191	0.3986
Prof	0.4283	0.4291
Constant	−2.6294**	0.8569

Note: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

In brief, we found that the review results were influenced by the previous research outcomes of applicants. Furthermore, reviewers tend to pay attention to journal impact factors, a measure that is easy to use as a proxy indicator for research quality. However, it is dangerous to use journal impact as a procedure in evaluation at the individual level. According to Seglen (1992, 1994, 1997), although articles are published in journals with relatively high expected citation rates, it does not necessarily imply that the articles themselves are good quality.

6.3 Result from second-round peer review

Half of the applications entered into the second peer review round, and nine out of these finally were granted funding support. As in the first round, we measured the cognitive distances between each application and the new group of reviewers in the second round. Figure 4 shows the result. Compared with the results of the first review round, a certain pattern can be observed between cognitive distances and final results.

Cognitive distances measured by ABC analysis were still quite large, like the results in the first round, mostly concentrated between 0.95 and 1. On the other hand, cognitive distances obtained by the A-T analysis were scattered from 0 to 1, and clearly there are only a few applicants with extremely short or long cognitive distance to their reviewers.

Unlike the results from the first round, it can be seen that six of nine winners have relatively short cognitive distances (A-T analysis) with their reviewers, while two of nine winners have high levels of cognitive distances (ABC analysis). Applicants located in the middle of the range had a small probability of success: only one application was granted. This result, on one hand, is consistent with statements from previous research that reviewers are predicted to be more likely to support applicants within short cognitive distances. On the other hand, our results show that reviewers also support applicants who have a long cognitive distance to the reviewers. In other words, results from the second round indicate that reviewers are

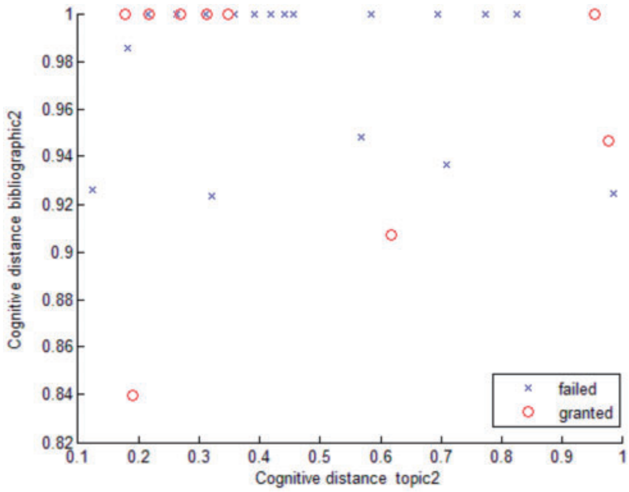

Figure 4. Cognitive distances between applicants and reviewers in second round.

likely to support applicants whose cognitive distance is either short or long, and that the applications in between have difficulties getting approval.

To summarize, in the first round, the impact counted as societal relevance and journal quality had a strong impact on the review results, but cognitive distance did not have a strong influence on the results. However, in the second round, an influence from the cognitive distance factor can be observed. The applicants with short or long cognitive distances had higher probabilities of getting granted.

7. Discussion

This article focuses on exploring the role of cognitive distance in the peer review process and measuring cognitive distances between applicants and referees. One motivation behind our work is lack of theoretical and empirical research on cognitive distance and peer review. The evidence here shows that cognitive distance was a neglected dimension when the SSF selected referees for reviewing the research proposals. In peer review procedures, academic status plays an important role. However, in the normal case, cognitive bias in the peer review process is not taken into consideration by those who are in charge for selecting reviewers.

Our results to some extent complement the earlier research on the role of cognitive bias. Previous research (Mahoney 1977; Travis and Collins 1991; Sandström 2009b; Sandström et al. 2010) focused on exploring the influence of cognitive similarity to the results of peer review, but ignored the impact caused by extremely low similarity of research content. Our results from the second review stage show that reviewers are likely to support the applicants who hold either short or long distances with them, which implies reviewers are more likely to approve applications with which they are familiar, and, actually, the same applies for applications with which they are relatively unfamiliar.

From the perspective of knowledge management, Nooteboom and colleagues (1999, 2000, 2007) have proposed and empirically confirmed that inverted U-shaped relations exist between cognitive distance and absorptive capacity. This implies that a too small or a too large cognitive distance has a negative effect on knowledge absorption. Therefore, there is an 'optimal cognitive distance' in a learning process. Borrowing the concept of 'optimal cognitive distance' and applying it to the peer review process, each application could be regarded as novel knowledge to its corresponding reviewer; meanwhile the absorptive capacity of the reviewer is dependent on the cognitive distance to the application. Figure 6 shows in the general case that as cognitive distance increases, absorptive capacity decreases whereas the novelty of knowledge increases. It implies that if the cognitive distance is small, a reviewer would be familiar with the research topic presented in the application; otherwise, the reviewer may have difficulties to completely understand or learn from the application. However, when the reviewer is very familiar with the research topic, i.e. active in the same research line, that might also cause cognitive bias in the peer review process, since the reviewer may take the view that the application is lacking innovativeness. Thus, in order to avoid cognitive bias and keep fairness, we assume that reviewers should have optimal cognitive distances with applicants. Optimal cognitive distance is supposed to be a certain range around the intersection point of absorptive capacity and novelty, denoted by O in Fig. 5. Meanwhile, the type of cognitive biases caused by short or long cognitive distance is also summarized in the figure.

We have noticed that current research and policy discussion in the area of computer science focuses on selecting reviewers conducting (very) similar research to the applicants or contributors. Many algorithms, based not only on references but also on research content, are applied to calculate the research similarity. However, the research community of computer scientists has ignored the bias caused by cognitive similarity. Hence, to avoid cognitive bias and keep fairness in the peer review process, we would like to suggest that funding agencies should avoid selecting reviewers whose cognitive distance to their applicants is either too large or too small.

Coming back to our case study, the results obtained especially from the second review stage, to some extent, contradict the pattern that would be expected from the theory of 'optimal cognitive distance'. To be more specific, our case shows that reviewers are much more likely to approve applications that they are familiar with, which is not consistent with the proposed theory of optimal distance. In our opinion, this can be seen from at least two perspectives.

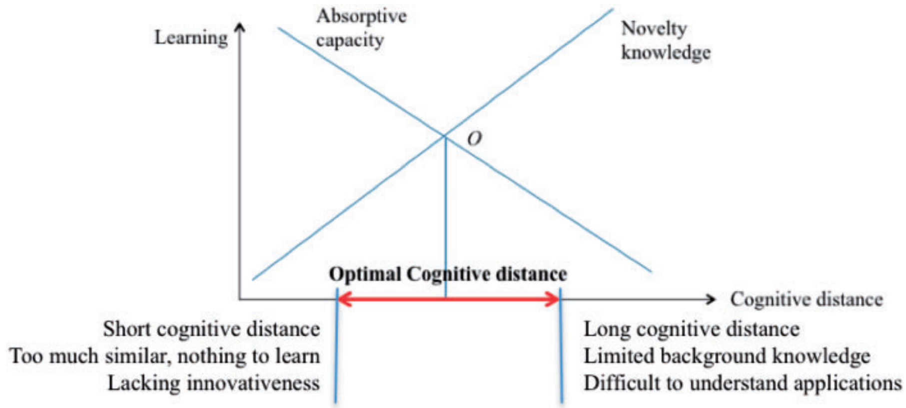

Figure 5. An explanation of optimal cognitive distance in peer review. (Note: The figure is revised based on Nooteboom et al. 2007.).

First, we need to emphasize that because cognitive distances obtained by ABC analysis in both peer review rounds were all quite large, our conclusions drawn depended more on the analysis of cognitive distances obtained by the A-T model. The reason of long distances obtained by ABC might be that most of the reviewers in the first review round were from the practical-industrial field and they did not have many publications. Besides, reviewers were probably not active in the same time frame as the applicants, since the number of publications and cited half-life of *infectious diseases* is 7,253 and 5.2, indicating the rapid renewal speed of this area. Thus, there might be very little overlap of the references used by reviewers and applicants (see Fig. 6). More important, it also implies although reviewers once worked in the same research area with applicants (short cognitive distance obtained by A-T model), they are not quite active now (long cognitive distance obtained by ABC model). In this instance, applications might still be quite new and innovative to the reviews. This is also an explanation why we insisted on measuring cognitive distance from both perspectives.

Second, the discrepancy could also be explained by factors other than cognitive distance. For instance, reviewers are less strict in the evaluation of unfamiliar applications for which they are not experts or they are likely to support research that is similar to their own. Therefore, from this normative point of view, reviewers should be selected from the range of 'optimal cognitive distance' to avoid the potential biases. Meanwhile, it is undeniable that further research both theoretical and practical is still required to explain this phenomenon.

In addition, it is also necessary to discuss limitations of the specific case in our study. As mentioned by Bornmann (2008, 2011), there are two fundamental problems making generalization of the findings from research on peer review difficult. First, it is difficult to judge whether the applicants who receive unfavorable review results are negatively evaluated due to the potential bias, like cognitive bias, or due to their 'insufficient quality of the proposals or

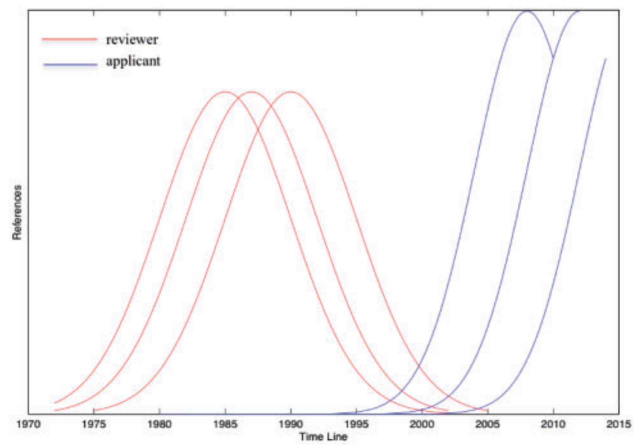

Figure 6. References used by reviewers and applicants.

manuscripts' (Bornmann 2008; see also Daniel 2004). Another limitation is that due to data access problems, lack of data makes the empirical research on fairness in the peer review process, such as research on cognitive bias, quite difficult. Our case study has the same problem because we have no information regarding the research proposals themselves. As a result, an assumption for this research is that all proposals should have the similar research quality or that the track record of applicants should count as the quality indicator. The latter perspective has been developed by several researchers (Wennerås and Wold 1997; Jayasinghe 2003; Bornmann 2011; Sandström and Hällsten 2008).

Meanwhile, lack of information of research proposals may cause another problem. There are cases when a researcher starts a new research trail or research line with a proposal to the research council. Obviously, in those cases the researcher does not have any papers in that trail and there will be no connection to reviewers with a specific bias for such a trail, although the trail will be opened by the researcher. However, due to the limited data, it is difficult to make further studies regarding this issue.

8. Conclusion(s)

The article explores the issue of cognitive bias in the peer review process. A major part of the article is an elaboration of the concept of cognitive distance in relation to the peer review processes. We show that there might be a theoretical connection between the concept of cognitive distance in the context of peer review and the research design, including the use of the A-T. Third, we illustrate a novel perspective to select reviewers, especially for the managers of research funding agencies, who as an actor might have large effects on cognitive development (Braun 1998). However, the analysis in this article was based on a relatively small number of cases, thus causing some statistical significance problems. Therefore, more empirical tests on the relations of cognitive distance and peer review are required before there is reason to implement new policies in these matters.

Acknowledgement

The research was made possible by grants from the Chinese Ministry of Education and the Swedish Foundation for Humanities and Social Sciences (R.J). The authors would like to thank Peter van den Besselaar of VU Univeristy Amsterdam, Ulf Heyman of Uppsala University for their comments on this article. We also thank Per Ahlgren of Stockholm University, Ludo Waltman, and other colleagues at the Center for Science and Technology Studies (CWTS) of Leiden University for their helpful comments on earlier versions. We would like to thank the editor and two anonymous reviewers for their valuable comments to improve the manuscript.

References

Ahlgren, P. and Colliander, C. (2009) 'Document-document Similarity Approaches and Science Mapping: Experimental Comparison of Five Approaches', *Journal of informetrics*, 3: 49–63.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003) 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, 3: 993–1022.

Bornmann, L. (2008) 'Scientific Peer Review: An Analysis of the Peer Review Process from the Perspective of Sociology of Science Theories', *Human Architecture: Journal of the Sociology of Self-knowledge*, 6/2: 23–38.

——. (2011) 'Scientific Peer Review', *Annual Review of Information Science and Technology*, 45/1: 197–245.

Bornmann, L. and Daniel, H. D. (2005) 'Selection of Research Fellowship Recipients by Committee Peer Review', *Scientometrics*, 63/2: 397–420.

Boyack, K. W., Klavans, R. and Börner, K. (2005) 'Mapping the Backbone of Science', *Scientometrics*, 64/3: 531–74.

Braun, D. (1998) 'The Role of Funding Agencies in the Cognitive Development of Science', *Research Policy*, 27: 807–21.

Caelleigh, A. S., Shea, J. A. and Penn, G. (2001) 'Selection and Qualities of Reviewers', *Academic Medicine*, 76/9: 914–5.

Cameron, A. C. and Trivedi, P. K. (2005) *Microeconometrics Methods and Applications*. New York: Cambridge University Press.

Cantner, U., Meder, A. and ter Wal, A. L. J. (2010) 'Innovator Networks and Regional Knowledge Base', *Technovation*, 30/9–10: 496–507.

Chubin, D. E. and Hackett, E. J. (1990) *'Peerless Science: Peer Review and US Science Policy'*. Albany: State University of New York Press.

Dangelico, R. M., Garavelli, A. C. and Petruzzelli, A. M. (2010) 'A System Dynamics Model to Analyze Technology Districts' Evolution in a Knowledge-based Perspective', *Technovation*, 30/2: 142–53.

Daniel, H. D. (2004) 'Guardians of Science: Fairness and Reliability of Peer Review', <http://onlinelibrary.wiley.com/book/10.1002/3527602208>.

Daud, A. (2012) 'Using Time Topic Modeling for Semantic-based Dynamic Research Interest Finding', *Knowledge-based Systems*, 26: 154–63.

Glänzel, W. and Czerwon, H. J. (1996) 'A New Methodological Approach to Bibliographic Coupling and its Application to the National, Regional and Institutional Level', *Scientometrics*, 37/2: 195–221.

Hautala, J. (2011) 'Cognitive Proximity in International Research Groups', *Journal of Knowledge Management*, 15/4: 601–24.

Hirsch, J. E. (2005) 'An Index to Quantify an Individual's Scientific Research Output', *Proceedings of National Academy of Sciences of the United States of America*, 102/46: 16569–72.

Hopcroft, J., Khan, O., Kulis, B. and Selma, B. (2004) 'Tracking Evolving Communities in Large Linked Networks', *Proceedings of the National Academy of Science*, 101: 5249–53.

Jayasinghe, U. W. (2003) *Peer Review in the Assessment and Funding of Research by the Australian Research Council*. Australia: University of Western Sydney, Greater Western Sydney.

Kessler, M. M. (1963) 'Bibliographic Coupling between Scientific Papers', *American Documentation*, 14/1: 10–25.

Ma, R. (2012) 'Author Bibliographic Coupling Analysis: A Test Based on a Chinese Academic Database', *Journal of Informetrics*, 6: 532–42.

Mahoney, M. (1977) 'Publication Prejudices: An Experimental Study of Confirmatory Bias in Peer Review System', *Cognitive Therapy and Research*, 1/2: 161–75.

McCullough, J. (1989) 'First Comprehensive Survey of NSF Applicants Focuses on their Concerns about Proposal Review', *Science, Technology, and Human Values*, 14/1: 78–88.

Mimno, D. and McCallum, A. (2007) 'Expertise Modeling for Matching Papers with Reviewer', *KDD'07 Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 500–9. New York: ACM Digital Library.

Nooteboom, B. (1999) *Interfirm Alliances: Analysis and Design*. London: Routledge.

——. (2000) 'Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance', *Journal of Management and Governance*, 4: 69–92.

Nooteboom, B., van Haverbeke, W., Duyster, G., Gilsing, V. and Oord, Al. (2007) 'Optimal Cognitive Distance and Absorptive Capacity', *Research Policy*, 36: 1016–34.

Pouris, A. (2007) 'Peer Review in Scientifically Small Countries', *R&D Management*, 18/4: 333–40.

Rosen-Zvi, M., Griffiths, T., Steyvers, M. and Smyth, P. (2004) 'The Author-topic model for Authors and Documents', *UAI'*

04 Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, pp. 487–94. Virginia: AUAI Press Arlington.
Salton, G. and McGill, M. J. (1983) *Introduction to Modern Information Retrieval*. Auckland, New Zealand: McGraw-Hill.
Sandström, U. (2009a) 'Bibliometrics Evaluation of Research Programs: A Study of Scientific Quality', <http://www.diva-portal.org/smash/get/diva2:486508/FULLTEXT01.pdf>.
——. (2009b) 'Cognitive Bias in Peer Review: A New Approach', *In Proceedings of 12th International Conference on Scientometrics and Informetrics*, July 28–31, Rio de Janeiro, Brazil, pp. 1–5.
Sandström, U. and Hällsten, M. (2008) 'Persistent Nepotism in Peer-Review', *Scientometrics*, 74/2: 175–89.
Sandström, U., Wold, A., Jordansson, B., Ohlsson, B. and Smedberg, Å. (2010) 'Hans Excellens: om miljardsatsningarna på starka forskningsmiljöer', <http://forskningspolitik.se/DataFile.asp?FileID = 194>.
Seglen, P. O. (1992) 'The Skewness of Science', *Journal of the American society for Information Science and Technology*, 43/9: 628–38.
——. (1994) 'Causal Relationship between Article Citedness and Journal Impact', *Journal of the American Society for Information Science and Technology*, 45/1: 1–11.
Seglen, P.O. (1997) 'Why the Impact Factor of Journals should not be Used for Evaluating Research', *British Medical Journal*, 314: 498–513.
Shitaba, N., kajikawa, Y., Takeda, Y. and Matsushima, K. (2009) 'Comparative Study on Methods of Detecting Research Fronts Using Different Types of Citation', *Journal of the American Society for Information Science and Technology*, 60/3: 571–80.
Sugimoto, C. R. and Cronin, B. (2013) 'Citation Gamesmanship: Testing for Evidence of Ego Bias in Peer Review', *Scientometrics*, 95/3: 851–62.
Swedish Foundation for Strategic Research (SSF), Framework grants, Infection biology. (2013), <http://www.stratresearch.se/en/Ongoing-Research1/Framework-grants/Infection-biology/>.
Travis, G. D. L. and Collins, H. M. (1991) 'New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System', *Science, Technology and Human Values*, 16/3: 322–41.
Van Raan, A. F. J. (2005) 'Measuring science: Capita selecta of current main issues'. In: Moed, H. F., Glänzel, W. and Schmoch, U. (eds) *Handbook of Quantitative Science and Technology Research*. New York: Springer.
Waltman, L., Jan van Eck, N., van Leeuwen, T. N., Visser, M. S. and van Raan, A. F. J. (2011) 'Towards A New Crown Indicator: An Empirical Analysis', *Scientometrics*, 87/3: 467–81.
Wennerås, C. and Wold, A. (1997) 'Nepotism and Sexism in Peer-Review', *Nature*, 387: 341–3.
Wesseley, S. (1998) 'Peer Review of Grant Applications: What Do We Know?', *Lancet*, 352: 301–5.
Whitley, R. (2000) *The Intellectual and Social Organization of the Science*. 2nd edn. Oxford: Oxford University Press.
Wuyts, S., Colombo, M. G., Dutta, S. and Nooteboom, B. (2005) 'Empirical Tests of Optimal Cognitive Distance', *Journal of Economic Behavior and Organization*, 26/6: 813–40.
Zhao, D. and Strotmann, A. (2008a) 'Author Bibliographic Coupling: Another Approach to Citation-based Author Knowledge Network Analysis', *Proceedings of the American Society for Information Science and Technology*, 45/1: 1–10.
——. (2008b) 'Evolution of Research Activities and Intellectual Influences in Information Science 1996–2005: Introducing Author Bibliographic-Coupling Analysis', *Journal of the American Society for Information Science and Technology*, 59/13: 2070–86.

Appendix

Table A.1. Measurement of the variables

Variable name	Measurement
Forward	Dummy variable set to one if the application entered to the second review round
Frac P	Sum of author fractionalized papers
NJCS	The impact of the journal set normalized in relation to its sub-fields (average = 1.00)
Top1%	The share of articles cited above the 99th percentile.
Vitality	Mean reference age normalized in relation to the sub-field set (average = 1.00, higher = younger)
Gender	Dummy variable set to one if applicant is female
Prof	Dummy variable set to one if applicant is professor

Note: detailed information of variables can be found in report of Sandström (2009a, P13).

Table A.2. Descriptive statistics of the variables

Variable name	Mean	Median	Standard deviation	Max	Min
Forward	0.48	0	0.50	1	0
Frac P	6.22	4.99	4.29	28.86	0.70
NJCS	1.03	0.97	0.32	2.54	0.48
Top1%	0.01	0	0.02	0.15	0
Vitality	1.02	1.02	0.09	1.41	0.76
Gender	0.16	0	0.37	1	0
Prof	0.67	1	0.47	1	0