圖書資訊學刊　第14卷第1期　（2016.6）　頁41-85
doi:10.6182/jlis.2016.14(1).041

同儕審查的起源、研究現況與展望

History, Research, and Challenges: A Systematic Analysis of Peer Review for Journals, Grants, and Faculty Appointments

黃慕萱[1]　嚴竹蓮[2]

Mu-Hsuan Huang[1], Chu-Lien Yen[2]

摘　要

同儕審查是學術界進行科學探索時所採用的一項自律機制，幾乎已制度化地納入學術組織的運作之中，並普遍獲得學界人士的支持。基本上同儕審查的正當性是基於學術社群成員之間的信賴與誠信，在各項學術活動中以不同的作業模式分配有限資源，包括學術文獻出版、研究計畫獎助、大學教職聘用與升遷，以及學術成就獎勵等。但是同儕審查的運作方式迄今未臻完善，除了出現效用、效率，以及信度等問題外，許多研究亦已證實存在多種評審者偏見，因此有愈來愈多的學者主張對同儕審查進行持續性地檢驗與監督，以提升評審作業的品質與公平性。本文透過文獻分析論述同儕審查在學術領域的應用與研究，首先說明同儕審查的定義、分類與優缺點，並以學術期刊稿件、獎助計畫，以及大學教職聘用／升遷之同儕審查為例，探討同儕審查的起源、發展與研究現況，最後分析同儕審查的國際合作及其與書目計量的關係。

關鍵字：同儕審查、書目計量

[1] 國立臺灣大學圖書資訊學系暨研究所
Department and Graduate Institute of Library and Information Science, National Taiwan University, Taipei, Taiwan

[2] 中華民國外交部資訊及電務處
Department of Archives, Information Management and Telecommunications, Ministry of Foreign Affairs, Taipei, Taiwan

* 通訊作者Corresponding Author: 嚴竹蓮Chu-Lien Yen, E-mail: clyen01@mofa.gov.tw

臺灣大學學術期刊資料庫

Abstract

Peer review is a self-regulation mechanism for scientific inquiry. Institutionalized and incorporated into the structure and operation of science, it has received considerable support in the academic setting. The legitimacy of peer review is based on trust and integrity. In various ways, it allocates scarce resources such as journal space, research funding, faculty recruitment, recognition, and rewards for academic achievements. But there are growing indications that peer review has yet to fulfill its potential functions, leading to negative assessments as to whether it is effective, efficient, or reliable. Many studies have found links between potential sources of bias and judgments in peer review and expressed reservations over the fairness of the process. It is, therefore, important that the peer review process should be subjected to serious scrutiny and regular evaluation that would lead to better quality and greater fairness. This study presents a systematic review of the empirical literature on peer review of journal manuscripts, grant applications, and faculty appointments and promotions. Historical and contextual information is provided as a basis for interpreting this review. Finally, the authors discuss international recommendations for good practice in peer review and the potential and problems of peer review and bibliometrics.

Keywords: Peer Review; Bibliometrics

Extended Abstract

1. Introduction

Peer review is a relatively new topic of research in the contemporary academic domain. It was initially used as an aid to scientific inquiry, to allow scientists to assess the quality of specific pieces of scientific research, in order to generate trustworthy and valid knowledge. Many scholars trace the origins of peer review back to the founding in 1665 of the journal *Philosophical Transactions* by the Royal Society of London, England. At that time the content of each issue would be reviewed by members of the Society before publication, and this has been seen as the fountainhead of journal peer review (Burnham, 1990; Kronick, 1990; Lock, 1985; Rennie, 2003; Spier, 2002a). Later, the scope of application of peer review gradually expanded, and it became one of the main mechanisms for the allocation of limited funding resources and for faculty appointments and promotions in higher education (Bornmann, 2011a; Frodeman, Holbrook, & Mitcham, 2012; Kronick, 1990; Langfeldt & Kyvik, 2011).

Note. This extended English abstract is supplied by the authors.

To cite this article in APA format: Huang, M.-H., & Yen, C.-L. (2016). History, research, and challenges: A systematic analysis of peer review for journals, grants, and faculty appointments. *Journal of Library and Information Studies, 14*(1), 41-85. doi: 10.6182/jlis.2016.14(1).041 [Text in Chinese].

To cite this article in Chicago format: Mu-Hsuan Huang and Chu-Lien Yen. "History, research, and challenges: A systematic analysis of peer review for journals, grants, and faculty appointments." *Journal of Library and Information Studies* 14, no. 1 (2016): 41-85. doi: 10.6182/jlis.2016.14(1).041 [Text in Chinese].

臺灣大學學術期刊資料庫

For more than three centuries peer review has been valued in academia, but there have also been many voices raised in criticism: numerous scholarly articles have questioned the true effectiveness of peer review, have criticized it as costly and time-consuming, and have identified signs of unfairness in the evaluation process (Abdoul, Perrey, Amiel, et al., 2012; Bornmann, 2011a; Fang, 2011; Lee, Sugimoto, Zhang, & Cronin, 2013; Rennie, 2003; Sandström & Hällsten, 2007; Wenneras & Wold, 1997; Wood & Wesseley, 2003). Nevertheless, most scholars continue to see the need to maintain this mechanism, because the possible alternatives are yet more controversial. Some researchers have even drawn a comparison with Churchill's description of democracy as the worst form of government, except all the other forms that have been tried (Harley, Acord, Earl-Novell, Lawrence, & King, 2010; Ismail, Farrands, & Wooding, 2009; Kostoff, 2004; Rennie, 1986; Sieber, 2006).

While academia has long recognized imperfections in the operation of peer review, it was not until around the 1980s that the need for rational scrutiny of peer review gradually began to gain attention. To date, empirical research on peer review has mostly concentrated on the review of journal manuscripts, followed next by that of grant applications (Bornmann, 2011a; Demicheli & Di Pietrantonj, 2007; Jefferson, Rudin, Brodney-Folse, & Davidoff, 2007; Weller, 2002; Wood & Wessely, 2003); there has been little research into peer review for faculty appointments (Miller, 1978; Weiser, 2012). In a review article on journal and grant peer review during the preceding decade, Bornmann (2011a) found that the bulk of such research examined issues of trustworthiness and fairness; there was relatively little examination of the predictive validity of peer review.

In addition to scholars' approaching peer review as a branch of academic knowledge, and constructing values and orientations for related scientific research, organizations that make use of peer review have also been actively engaging in international exchange, including publishing basic operational standards or practice guides for peer review, in order to improve the quality of evaluation and enhance its credibility. However, as yet there has been little research that confirms the validity of peer review, so that there is a lack of firm correlation between review outcomes and levels of contribution to science. Some scholars have suggested that peer review should be reinforced by the use of bibliometrics, but how these two assessment methods can be integrated is an issue that still awaits academic debate and research.

2. Definition, Categories, and Pros and Cons of Peer Review

"Social psychology conceptualizes the peer review process as a social judgment process of individuals in a small group (for example, one or more reviewers and one or more editors of a disciplinary in-group in manuscript reviewing)" (Bornmann, 2011a, p. 200). Some researchers have likened peer review to the technique of "grading the grain" in agricultural produce markets. However, the grading done by peer review is not a single procedure, but a set of flexible and adjustable mechanisms. Each organization that uses peer review has a different way of operating, and there may even be differences among different departments within the same organization (Chubin & Hackett, 2003; Cole & Cole, 1973; Kostoff,

臺灣大學學術期刊資料庫

2004; Research Councils UK [RCUK], 2006; Research Information Network [RIN], 2010).

The scope of application of peer review in academia is very broad, and peer review has been subcategorized in different ways, mainly based on the target of evaluation (Harley & Acord, 2011; Parliamentary Office of Science and Technology [POST], 2002; RIN, 2010). In this study, we divide peer review into three major types: peer review of publications, peer review for funding purposes, and peer review of cumulative achievement.

In theory, the advantages of peer review appear almost self-evident: Scientists' peers are assumed to be naturally the people most able to put forward correct opinions and suggestions on their research (Chubin & Hackett, 1990). Proponents of peer review believe that its effects are at least better than those achieved by self-restraint on the part of scientists, and that it can provide constructive rational assessment in order to improve the quality of research (Bornmann & Daniel, 2005; Nickerson, 2005). Many studies have indicated broad acceptance of and high levels of support for peer review in academic circles (Bertout & Schneider, 2004; Boden et al., 1990; Fletcher & Fletcher, 2003; Gibson, Spong, Simonsen, Martin, & Scott, 2008; POST, 2002; Royal Society, 1995; Ware & Monkman, 2008). However, peer review has also attracted much criticism, for example that the review process is time-consuming and costly; that the effectiveness of peer review is not readily demonstrable; that there is a low degree of reliability between evaluators; and that review outcomes are unfair, fail to encourage innovation, and are unfavorable to young scientists.

Today, peer review is a mechanism of self-regulation in academic communities, and has been closely incorporated into the operations of academic organizations. Yet many scholars believe that peer review has inherent limitations.

In this study we present eight major characteristics of peer review: the difficulty of defining peers; the conflicted role of the peer; flexibility in the interpretation of review criteria; the need for both originality and skepticism to assure the quality of science; compromise between multiple goals; decision-making by a small number of people behind closed doors; challenges arising from the demands for both academic autonomy and social responsibility; and the need for balance between trust and oversight. These are all issues that need to be confronted when organizing peer review activities.

3. Origins and Development of Peer Review

Peer review is a core element of the scientific domain. It plays a crucial role in the publication of journal articles, in the allocation of funding resources, and in the recruitment and career progression of university faculty members.

However, while the history of journal peer review can be traced back to the publication of *Philosophical Transactions* by the Royal Society of London from the 17th century onward, grant peer review and faculty appointment peer review are mechanisms that developed independently in the 19th and 20th centuries in response to practical operational needs. Today, peer review mechanisms are widely used by research funding bodies and institutes of higher education (Bornmann, 2011a; Weiser, 2012).

The modern development of peer review has been influenced by bibliometrics. The impact

臺灣大學學術期刊資料庫

factor proposed by Garfield and Sher (1963) has given concrete expression to the validity of journal peer review. In the late 20th century, various bibliometric indicators, in particular citation analysis, gradually became important points of reference for evaluating the research performance of individuals and institutions, so that through the intermediary of bibliometrics, journal peer review once again indirectly became an important factor influencing funding allocation and faculty appointments (Harley & Acord, 2011; Harley et al., 2010; van Arensbergen, van der Weijden, & van den Besselaar, 2014a). In addition, for journal readers in the age of the information explosion, journal peer review's function of quality control over the scientific literature makes it a reliable guarantee of reading quality.

Through more than 300 years of development, journal peer review has become the gatekeeper of scientific knowledge, and is greatly valued by the scientific community. But there are also voices of criticism and calls for reform, and with the development of electronic journal publishing and continued improvements in its quality, along with changing reading habits, many scholars believe that in the not too distant future, revolutionary changes may take place in journal peer review.

The development of grant peer review, on the other hand, is closely related to expansion in the functions of government. Particularly in the 20th century, various countries began to devote funds to supporting scientific research, and peer review became one of the major mechanisms by which governments allocate funding (Burnham, 1990; Frodeman et al., 2012). At present, important issues in grant peer review include the evaluation of indicators of social impact and of innovation, with both presenting difficulties in selecting peers as well as in defining review criteria (Bell, Shaw, & Boaz, 2011; van der Meulen & Rip, 2000).

As for faculty appointments peer review, in an environment of expansion in university education and the systematization of faculty recruitment, this type of peer review has gradually become the arbiter of scholars' access to campus teaching positions (Snodgrass, 2006; Weiser, 2012). In faculty appointments peer review, there is a general issue of overemphasizing research (Fairweather, 2005; Greenbank, 2006; Kreber, 2002; Pratt, 1997) and undervaluing teaching. In recent years many countries have taken note of this phenomenon and have actively adjusted their higher education policies to encourage high-quality teaching, in order to promote a better balance between research and teaching (Hénard, 2010).

4. Current Status of Peer Review Research

Peer review is a mechanism for the allocation of scarce resources in academia, yet its processes are confidential in nature and there are few clues to them in the public domain. It was not until around the 1980s that peer review began to be subjected to a greater degree of rational scrutiny (Chubin & Hackett, 1990; Rennie, 2003; Weller, 2002). To date, research into peer review has mainly concentrated on peer review of journal manuscripts and grant applications; there has been very little investigation of peer review for faculty appointments and promotion in higher education.

However, numerous researchers have observed that in the great majority of the related literature the methodology is weak and the

臺灣大學學術期刊資料庫

causal reasoning tenuous, and that the great diversity of research results makes it difficult to reach general conclusions. They suggest that such studies can be strengthened by using meta-analysis or experimental research (Bornmann, 2011a; Bornmann, Nast, & Daniel, 2008; De Vries, Marschall, & Stein, 2009; Demicheli & Di Pietrantonj, 2007; Jefferson et al., 2007). Marsh, Jayasinghe and Bond (2011) suggest that in addition to meta-analysis of secondary data from the literature, there is also a need for large-scale research on primary data from the peer review process itself.

The development of journal peer review is being impacted by the Internet and electronic publishing. Some journal publishers and research institutions have taken the initiative to trial new forms of evaluation, such as pre-publication peer review, post-publication peer review, and open peer review, in the hope of establishing as soon as possible the most appropriate mode of evaluation for journal manuscripts in the digital age (Borgman, 2007; Bornmann & Daniel, 2010; Bornmann, Marx, Schier, Thor, & Daniel, 2010; Ford, 2013; Odlyzko, 1996; van Rooyen, Delamothe, & Evans, 2010).

In the field of research into grant peer review, with the trends toward open access to government information and toward government performance management, there has been much progress in recent years. More and more government funding agencies are willing to make data on the evaluation process accessible to researchers, so that scholars are no longer limited to analyzing only evaluation outcomes, but can also analyze and compare different stages of the peer review process in detail (Bornmann, Leydesdorff, & van den Besselaar, 2010).

Although there has been little research into faculty appointments peer review, more and more scholars are giving attention to its objectivity and fairness. ACUMEN (Academic Careers Understood through Measurement and Norms), funded under the European Commission's Seventh Framework Programme, integrates the three evaluation methods of peer review, bibliometrics, and webometrics to create a framework—the ACUMEN Portfolio—for the evaluation of individual academic performance, and the ACUMEN Consortium has published detailed practice guidelines for the use of organizations responsible for funding allocations and faculty appointments. In an ACUMEN Portfolio, the first part presents the candidate's own career narrative, while the following three parts present indicators of the candidate's expertise, output, and impact (ACUMEN Consortium, 2014; Tatum & Wouters, 2013).

5. Future Development of Peer Review

In the great majority of countries, the operating methods of peer review are not subject to statutory regulation, but are decided by the user organizations themselves. There may be explicit practice guidelines, or evaluation may be conducted according to established internal procedures; the methods employed by different institutions are highly diverse (Bornmann, 2011a; General Accounting Office, 1999; Organisation for Economic Co-operation and Development, 2011a; Rennie, 2003; Weiser, 2012; Weller, 2002; Wood & Wessely, 2003).

臺灣大學學術期刊資料庫

However, for many years peer review users in various countries have also actively engaged in international exchange, sharing their experience in order to improve the quality and credibility of peer review. Such international collaboration has already borne fruit in the areas of both journal peer review and grant peer review, and some scholars have also advocated international collaboration between institutes of higher education to jointly define basic evaluation indicators for faculty appointments peer review (Harley & Acord, 2011; Weiser, 2012).

In addition to conducting research into peer review, the academic world has also been actively seeking improvements and alternative solutions. In recent years, many innovative methods have emerged in journal peer review, which is showing diverse development. In the areas of grant peer review and faculty appointments peer review, the aspect that most often provokes debate is the competitive-cooperative relationship between peer review and bibliometrics.

Some scholars believe that the use of a broad and diverse range of bibliometric indicators is conducive to enhancing the rationality and transparency of peer review (Bornmann, 2011b, 2013a; van Raan, 2005). Geisler (2001) observed that the combined use of peer review and bibliometrics enables peer review decisions to be no longer merely the subjective opinions of the evaluators, but introduces objective quantitative indicators. However, how to appropriately incorporate the quantitative data of bibliometrics into different and diverse peer review mechanisms presents a major challenge to the academic world.

6. Conclusion

Among all human activities, scientific research may be one of those most subject to benchmarking and evaluation (Laloë & Mosseri, 2009), both in order to assure the quality of research and as a mechanism for allocating scarce academic resources. But peer review has inherent limitations. On the one hand, as a human activity it may easily be affected by human weaknesses or biases; and on the other hand reviewers wield immense power, but emphasize confidentiality (Geisler, 2000; Wenneras & Wold, 1997; Ziman, 2000), and so attract criticism of "black box" operations. Some scholars have even asserted that the only reason for continuing to use peer review is the lack of a better alternative (Kostoff, 2004; Rennie, 1986; Sieber, 2006).

From the perspective of the spirit of academic autonomy, the legitimacy of peer review is based in mutual trust and honesty between members of the academic community. The British Academy's (2007) report on peer review points out that when reviewing documents, reviewers may come into contact with their peers' original data sets, new empirical results, or innovative conceptual frameworks. In the business world, such information would be regarded as trade secrets, but in the academic domain evaluators' act as part of the scientific regulatory system, and their ultimate goal is to rapidly and effectively disseminate research results. Therefore evaluators' integrity is of tremendous importance. However, numerous studies have demonstrated that various biases exist among evaluators, that there is too little consistency among evaluators, and that the effectiveness of evaluation is difficult to demonstrate. Thus there is a need to improve

臺灣大學學術期刊資料庫

the fairness, rationality, and transparency of peer review.

Today, peer review is widely valued in academia, and has become the main referee of various kinds of academic activity. However, academic disciplines are becoming increasingly specialized and complex, and the research population is growing ever larger, yet there has been no great increase in academic resources.

In these circumstances, competition will grow increasingly intense, and thus it seems likely that research into peer review will attract ever greater attention. This growing competition is evidenced by the facts that today acceptance rates for manuscripts submitted to many leading peer-reviewed journals are in single figures, and that rates of funding allocation by government and private-sector funding organizations in various countries are also trending downward year by year (National Institutes of Health, 2013; National Science Foundation, 2011; Powell, 2010; RCUK, 2006). Moreover, there has also been criticism regarding the fairness of evaluation methods used for faculty appointments and promotion in higher education. Therefore in the future peer review practice will inevitably face increasing questions and challenges from the outside world, and establishing a mechanism for continual supervision, scrutiny and improvement should be a direction for concerted effort within academia (Bornmann & Daniel, 2008; Callaham, 2003; De Vries et al., 2009; Gluckman, 2012; Godlee & Jefferson, 2003; Henly & Dougherty, 2009; Hojat, Gonnella, & Caelleigh 2003; Langfeldt, 2001; Rennie, 2003).

壹、前言

「同儕審查」（peer review）是當代學術領域中一個比較新的研究主題，它原本用以輔助科學探索，容許科學家運用它來評判某一特定科學研究的品質，以產生具有信度與效度的知識。之後同儕審查的應用範圍逐漸擴大，成為學術界分配有限資源的主要機制之一，在期刊文獻出版、獎助資源分配，以及大學教職的聘用與升遷上都扮演著關鍵角色（Bornmann, 2011b; Frodeman, Holbrook, & Mitcham, 2012; Kronick, 1990; Langfeldt & Kyvik, 2011）。近半個世紀以來，同儕審查的效用（effectiveness）、效率（efficiency）、信度（reliability），以及公平性（fairness）等受到多方質疑，有關同儕審查的研究也因此日益受到重視，研究的廣度與深度亦逐漸增加，而成為一個獨立且益形重要的學門。

同儕審查相應的英文同義詞除了peer review之外，還包括peer advice、peer evaluation、peer judgement、peer censorship、merit review，以及refereeing等（Chubin & Hackett, 1990）。中文也有不同的說法，例如同儕審查、同儕互評、同儕評論、同儕評鑑、同儕評估、同行評議、同行評審，以及同行評閱等，本研究將採用「同儕審查」做為前述中、英文含義的通用詞彙。

許多學者認為同儕審查可上溯至1665年英國倫敦皇家學會（Royal Society of London）創辦的《哲學學報》（*Philosophical Transactions*），當時每期學報在出刊前，均須經由學會會員審查內容，而被視為期刊同儕

臺灣大學學術期刊資料庫

審查的濫觴（Burnham, 1990; Kronick, 1990; Lock, 1985; Rennie, 2003; Spier, 2002a）。獎助同儕審查的發展則與政府的職能擴張息息相關，尤其在20世紀，各國政府開始編列預算支持科學研究，同儕審查就成為政府獎助經費分配的重要機制之一（Burnham, 1990; Frodeman et al., 2012）。至於大學教職聘用／升遷同儕審查是在大學教育普及化與教職聘用制度化的環境下，逐漸成為學者進入校園任教的仲裁者（Snodgrass, 2006; Weiser, 2012）。到了20世紀後期，英美等國政府強調科學管理，積極主導國家科學發展方向，同儕審查遂進一步成為科技政策制定及教研機構評鑑的重要工具，而運作方式與評審標準也更為多元（Frodeman & Briggle, 2012; Guston, 2003; Organisation for Economic Co-operation and Development [OECD], 2011a, 2011b; Whitley & Gläser, 2007）。

300多年來，同儕審查受到學術界的重視，日益擴大應用範圍，但是批評聲浪亦眾，許多文獻質疑同儕審查的真實效用、批評其過程耗錢費時，並且發現評審不公的現象，包括機構（institutional）偏見、交情（cronyism）偏見、年齡偏見、性別偏見、國籍偏見、非英語母語偏見、保守（conservative）偏見、學派偏見、學術產出（productivism）偏見，以及審察先後順序偏見等（Abdoul, Perrey, Amiel, et al., 2012; Bornmann, 2011b; Fang, 2011; Lee, Sugimoto, Zhang, & Cronin, 2013; Rennie, 2003; Sandström & Hällsten, 2007; Wenneras & Wold, 1997; Wood & Wesseley, 2003）。儘管如此，大部分學者依然認為有必要維持這個機制，因為其他替代方案的爭議性更大；有些學者甚至將同儕審查類比為自由世界的民主制度，是不好機制中的最佳選項（Harley, Acord, Earl-Novell, Lawrence, & King, 2010; Ismail, Farrands, & Wooding, 2009; Kostoff, 2004; Rennie, 1986; Sieber, 2006）。

雖然學術界早已認知到同儕審查作業的不完美，但是直到1980年代前後，同儕審查的理性檢驗才逐漸受到重視，1989年期刊同儕審查開始定期舉辦研討會，也帶動了一股研究風潮（Chubin & Hackett, 1990; Rennie, 2003; Weller, 2002）。目前同儕審查的實徵研究大多以期刊稿件為主，獎助計畫次之（Bornmann, 2011b; Demicheli & Di Pietrantonj, 2007; Jefferson, Rudin, Brodney-Folse, & Davidoff, 2007; Weller, 2002; Wood & Wesseley, 2003），大學教職聘用／升遷同儕審查的研究不多（Miller, 1978; Weiser, 2012）。Bornmann（2011b）回顧近十餘年的期刊及獎助同儕審查實徵研究，其中大多數為信度及公平性議題，預期效度的研究較少。

除了學術界將同儕審查視為一門學術知識，為其建構科學研究的價值與方向外，同儕審查的主事機構也積極進行國際交流，合作出版同儕審查的基本作業規範或操作指南，以強化評審品質並提升公信力。此外在網路科技發達之後，期刊同儕審查也出現許多創新作法，例如預印（pre-publication）同儕審查、出版後（post-publication）同儕

臺灣大學學術期刊資料庫

審查，以及公開（open）同儕審查等，目的都在強化評審作業的公平性與透明性。不過直到今天，同儕審查的效度依然少有研究證實，使得評審結果與科學貢獻度之間缺乏堅實的關聯性；有些學者建議以書目計量加以補強，但是如何整合這兩種評鑑方法，仍是學術界有待討論與研究的議題。

貳、同儕審查的定義、分類與優缺點

一、同儕審查的定義與分類

社會心理學家將同儕審查的過程概念化，稱之為「在小團體中，針對個人進行社會評價的過程，例如期刊編輯或評審者審查稿件（Bornmann, 2011b, p. 200）。」有些學者將同儕審查類比為農產市場的穀物分級技術，不過同儕審查的分級並非單一程序，而是一組有彈性且可調整的機制，各個主事機構有不同的作業方式，甚至同一機構的不同單位亦可能有所差異（Chubin & Hackett, 2003; Cole & Cole, 1973; Kostoff, 2004; Research Councils UK [RCUK], 2006; Research Information Network [RIN], 2010）。

以下列舉兩則學術界經常引用的同儕審查定義，一者就運作層面來看：「Michael Gibbons與Luke Georghiou認為同儕審查是某一特定領域的專家，對於相同或相近領域之其他科學家的研究作品，進行科學特定面向（如研究品質）的專業價值判斷。其前提係基於評審專家必須對領域的認知發展、研究方向，以及研究社群有充足的知識（OECD, 2011a, p. 1）。」一者強調目的與應用：「同儕審查是一套科學研究的品質評價機制，科學界利用同儕審查來確認研究程序的正確性及推論的合理性，並根據評審結果分配有限資源，例如期刊版面、獎助名額、學術聲望與特殊榮譽等（Chubin & Hackett, 1990, pp. 1-2）。」

同儕審查在學術界的應用範圍甚廣，分類方式亦不相同。英國人文社會科學院（The British Academy）將同儕審查分為兩個層級，第一級同儕審查是直接針對科學家的研究產出進行評審，例如期刊同儕審查及獎助同儕審查等。而若只是在評審過程中參考第一級同儕審查的結果，則屬於第二級同儕審查，例如高等教育機構評鑑或世界大學評鑑等，近年來第二級同儕審查的應用已有益形增加的趨勢（The British Academy, 2007）。Frodeman等人（2012）則依據評審的用途將同儕審查區分為三類，其一為預期性用途，期刊同儕審查及獎助同儕審查均屬此類，評審者不但要審查期刊稿件或獎助計畫的科學品質，還須顧及評審結果的預期效度；其二為反省性用途，主要檢驗學術領域所採行的某種制度是否達到預訂目標，例如期刊品質評鑑或學術機構評鑑等；其三為綜合性用途，結合預期性及反省性兩種用途，例如大學教職的聘用／升遷同儕審查，評審者除了回溯受評者的過去成就，也必須預測其未來的表現。

臺灣大學學術期刊資料庫

比較普遍採用的分類法是以評審標的物為基準，研究者綜整為三類：(一)出版品同儕審查：包括文獻類出版品（如期刊文獻及專書等）及非文獻類出版品（如紀錄片、資料庫、網站，以及電腦軟體等）的同儕審查；(二)獎助同儕審查：包括獎助計畫及獎學金的同儕審查；(三)成就同儕審查：評鑑個人、團隊、部門或機構的學術表現，並依據評審結果分配資源或獎勵，包括大學教職聘用／升遷同儕審查、學術榮譽或獎項同儕審查，以及教育暨研究機構評鑑等（Harley & Acord, 2011; Parliamentary Office of Science and Technology [POST], 2002; RIN, 2010）。

二、同儕審查的優缺點

理論上同儕審查的優點似乎不言自明，科學家的同儕當然是最能夠對其研究提出精確看法與建議的人（Chubin & Hackett, 1990）。支持者認為同儕審查的效果至少比科學家的自我約束為佳，並且可提供建設性的理性評價以改進研究品質（Bornmann & Daniel, 2005; Nickerson, 2005）。而若就宏觀層面的科學發展來看，科學研究必須經過批判與檢驗，才得以降低錯誤，成為真正的科學知識（Popper, 1961; Ziman, 2000）。

許多調查研究顯示，學術界對於同儕審查廣泛接受且高度支持（Bertout & Schneider, 2004; Boden et al., 1990; Fletcher & Fletcher, 2003; Gibson, Spong, Simonsen, Martin, & Scott, 2008; POST, 2002; Royal Society, 1995; Ware & Monkman, 2008）。Geisler（2000）整體論述同儕審查的優點，包括提供稱職專家意見、減少低品質科研、管控科研質量、平衡不同科學觀點或思想學派、採取理性、有效與公平的過程、承擔科學發展的責任，以及協助科技資源分配與決策。Brown（2004）及Ware（2013）則指出期刊同儕審查的多項功能，例如文獻的註冊、保存與改進品質、學術知識的過濾與淨化、學術領域範圍的形塑與發展，以及評審者個人知識的提升等。

不過同儕審查也招致許多批評，研究者綜整如下：(一)信度：評審者之間的一致性過低，可能存有潛在偏見；(二)公平性：評審過程受到非科學價值因素的影響，例如受評者的特殊背景或評審者的個人偏見等，而導致評審不公的情況；(三)效度：評審的效度證實不易，也缺乏客觀指標；(四)效率：評審作業耗費時日且所費不貲；(五)創新性：評審結果傾向於保守主義，未能鼓勵創新；(六)累積性：評審方式比較有利於著名或資深學者，不利於缺乏學術成就的年輕科學家。其他的批評還包括評審過程不透明、評審者剽竊研究創意，以及研究內容偽造或抄襲等，都是經常被討論的問題（Abbott, 2008; Bornmann, 2011b; Bornmann & Daniel, 2005; Braben, 2004; Chubin & Hackett, 1990; Gillet, 1993; Gluckman, 2012; Horrobin, 1990, 1996; Langfeldt & Kyvik, 2011; Luukkonen, 2012; RCUK, 2007; Rip, 2000; Roy, 1985; Travis & Collins, 1991）。

臺灣大學學術期刊資料庫

總之，同儕審查是學術社群的自我規範機制，並已緊密地納入學術組織的運作之中，許多學者認為同儕審查有其先天限制，研究者綜整提出同儕審查的八大特質，這些都是在規劃同儕審查作業時必須面對的議題。

(一)「同儕」的定義不易：根據同儕審查的定義，學術領域是決定評審者的主要關鍵，但是學術領域的界線本有爭議（Price, 1963），而且即使是相同學門的子領域之間，其專業性也未必可以彼此互通。近年來科學專業性的發展益加多元，愈來愈多跨領域、超學科，以及團隊科學的研究，使得同儕的界定更為困難（The British Academy, 2007; Hicks & Katz, 1996; Holbrook, 2013a, 2013b; Lamont, 2009; Langfeldt, 2001, 2006; Wood & Wessely, 2003），個別領域專家在評審跨領域研究時，極有可能面臨專業知識不足的問題（Bornmann, 2011b; Frodeman et al., 2012; Holbrook, 2013a; Huutoniemi, Klein, Bruun, & Hukkinen, 2010; OECD, 2011a）。

(二)「同儕」的角色矛盾：同儕審查最顯著的特徵之一是評審的主客體有可能互換，今天擔任評審者的專家，明天可能在另一場合與現在的受評者直接競爭，這種既是裁判又是球員的情況，很難避免利益衝突（Abdoul, Perrey, Tubach, et al., 2012; Langfeldt, 2001）。尤其在某些研究者人數較少的高度專業化領域，或者小規模科研國家，這個問題就更加複雜（Chubin & Hackett, 1990; OECD, 2011a; Pouris, 1988）。此一潛在風險也促使愛爾蘭及以色列等國家，在進行學術相關評鑑時只邀請外國評審者參與（Gluckman, 2012）。

(三)「評審標準」的解讀彈性：以期刊同儕審查為例，評審者除須判斷稿件的重要性與合理性外，還須對整體研究提出支持或否決的建議，這些都是具有高度主觀性的議題，每個人的看法不盡相同。目前同儕審查的主事機構大都在事前訂定評審標準或原則，但是評審者在實際操作時，往往對於既定的標準有不同的解讀或權重，因而造成評審者之間一致性過低的現象，也引發各界對評審不公的質疑（Bornmann, 2011b; Chubin & Hackett, 1990, 2003; Cicchetti, 1997; Kostoff, 2004; Rennie, 2003; Weiser, 2012; Ziman, 2000）。

(四)「科學品質」的求同求異：Polanyi（1962）指出科學的專業標準包括信度、科學價值，以及原創性，前二者主要強調科學研究應有其一致性，但是原創性則是鼓勵異議與批判，此兩項特色在本質上就有矛盾之處。同儕審查作為科學研究的評審機制，自然也會面臨相同的難題，除了要評量出具有一致性的核心研究外，還要能挑選出各個領域中最具潛力的創新想法，此一困境由各領域學者對於研究前沿（research front）的低度共識即可得證（Cole, 2000; Wood & Wessely, 2003）。

臺灣大學學術期刊資料庫

(五) 「多重目標」的取捨與妥協：同儕審查作業所涉及之利益關係者眾多，主事機構往往訂有多重目標，但是各個目標之間卻經常存在彼此消長的關係。以獎助同儕審查為例，獎助機構的目標通常包括效率、效用、課責性、回應性、穩定性、合理性、創新性，以及公平性等，但是若強調了課責性就必須犧牲一些創新性、提升穩定性將會降低回應性，而堅持了公平性則可能要以效用做為代價（Chubin, 1994; Chubin & Hackett, 1990; Luukkonen, 2012; RCUK, 2006; Wood & Wessely, 2003）。

(六) 少數人決定的「閉門民主」：同儕審查經常被批評為黑箱作業，主要原因有二，其一、同儕審查是由少數人挑選的少數評審者進行審查，審查報告在本質上雖然只是評審者的個人意見，但是在實際運作時卻代表整體科學家的決定（Harnad, 1996; Perper, 1989; Ziman, 2000）；其二、同儕審查作業強調機密性，在過程不公開且不透明的情況下，容易受到評審者個人價值觀或自身利益的影響，而造成評審結果不公（Chubin & Hackett, 1990; The Higher Education Academy, 2009; Rennie, 2003）。

(七) 「自治與責任」的挑戰：20世紀末期以來，各國政府要求學術領域在科學自治外，也要回應民眾的需求，兼顧科學研究的社會責任。現今全球許多國家的政府獎助機構已直接或間接地將社會影響（social impact）或國家利益納入評審考量，以爭取民眾對政府科研預算的支持。為此同儕審查的自治性與獨立性正在弱化，而逐漸成為一種服務國家政策目標的機制，這種科學認識論與政治問題結合的用途，更加強化了同儕審查作業的矛盾與兩難（Bhattacharya, 2012; Bornmann, 2013b; Collins & Tabak, 2014; Cozzens, 1999; Frodeman et al., 2012; Holbrook & Frodeman, 2011; Holbrook & Hrotic, 2013; Kamenetzky, 2012; Rip, 2000）。

(八) 「信賴與監督」的平衡：同儕審查的正當性是基於學術社群成員之間的彼此信賴與誠信（Abdoul, Perrey, Tubach, et al., 2012; Giraudeau, Leyrat, Le Gouge, Léger, & Caille, 2011; Williamson, 2003），有些學者認為Merton（1942）的科學四大規範（普遍主義、公有主義、無私利性，以及有條理懷疑主義）足以制約科學家的評審行為（Frodeman et al., 2012; Heitman, 2002），但是許多研究已經證實評審者存有各種偏見，並呼籲建置同儕審查的檢驗與監督機制，以提升評審的公平性（Biagioli, 2002; Callaham, 2003; De Vries, Marschall, & Stein, 2009; Gluckman, 2012; Levy, 1984; Smith, 2003; van Rooyen, Black, & Godlee, 1999）。

臺灣大學學術期刊資料庫

參、同儕審查的起源與發展

同儕審查是學術領域的核心，在期刊稿件出版、獎助資源分配，以及大學教職的聘用與升遷上都扮演著關鍵角色。不過就同儕審查的起源來看，除了期刊同儕審查有其歷史成因，可溯源自17世紀英國皇家學會出版的《哲學學報》（Biagioli, 2002; Kronick, 1990; Lock, 1985; Zuckerman & Merton, 1971）。獎助計畫與大學教職聘用／升遷的同儕審查，則是在19至20世紀間，因應實際作業需要而獨立發展的機制，今天研究獎助機構及大學已廣泛採用同儕審查機制（Bornmann, 2011b; Weiser, 2012）。

一、期刊同儕審查的起源與發展

(一) 期刊同儕審查的早期史

1. 出版檢查與稿件品質控管

學術界採用的同儕審查機制究竟起於何時，學者看法不一。Kronick（1990）認為就廣義來看，同儕審查遠在人類開始探索與傳播新知時，就已經存在，因為同儕審查（無論發生於出版前或出版後）是建立共識的必要作法，也是科學知識成長的基礎。而大部分的學者則將同儕審查溯源至17世紀英國皇家學會出版的《哲學學報》，該刊於1665年創刊，由學會首任秘書Henry Oldenburg擔任編輯，每期學報出刊前必須先由學會會員進行內容審查，而被視為期刊同儕審查的濫觴（Burnham, 1990; Kronik, 1990; Lock, 1985; Rennie, 2003; Spier, 2002a; Zuckerman & Merton, 1971）。

Biagioli（2002）認為早期期刊同儕審查與政府的出版檢查制度相關，15至16世紀，印刷術在歐洲快速傳播，歐洲各專制政權為了控制言論，實施出版特許制度，規定任何文字內容均須經由政府核可才得以印刷販售。17世紀中葉，科學研究逐漸機構化，歐洲各地由皇家支持的科學學會興起，並授予出版特許以推廣科學新知，此一作法等同將科學文獻的內容檢查工作交由科學學會負責。整體來看，當時科學學會的期刊同儕審查屬於國家出版檢查體系的一環，各學會除了評審稿件的科學品質外，也同時進行內容檢查，以排除與政府當局不一致的論點（Frodeman et al., 2012; Zuckerman & Merton, 1971）。

到了18世紀初期，學術期刊出版市場逐漸成形，一方面因為歐洲各國科學學會積極出版優良期刊，以建立聲望並鞏固地位；另一方面是科學家努力發表期刊文獻，以取得進入科學社群的資格，因此期刊同儕審查的稿件品質控管功能逐漸超越了內容檢查。此一發展也促使期刊同儕審查在歐洲專制政體垮台之後，得以擺脫出版檢查幫手的負面形象，由政府出版檢查制度的執行者，轉身成為期刊稿件品質的管理者（Biagioli, 2002; Lock, 1985）。

然而直到19世紀末期，學術期刊並未立即全面採用同儕審查，主觀因素在於期刊編輯或出版者憂心編輯大權旁落，不願意借重外界專家協助審核文稿；客觀因素則是期刊的數量快速成長，稿件不足的情況甚為

臺灣大學學術期刊資料庫

嚴重，當時期刊編輯的首要任務是尋找好文章來填補版面，必要時才將稿件交由專家審查（Burnham, 1990; Rennie, 2003; Spier, 2002a）。

2. 學科專門化與稿件供需逆轉

期刊普遍採用同儕審查主要出現在20世紀，原因之一是期刊稿件供需情勢逆轉，隨著學科專門化以及科學家的人數愈來愈多，研究文獻快速成長，使得期刊稿件不足的情況消失了，編輯工作的最大難題不再是尋找稿件來填補版面，而是如何在眾多投稿中挑選出最佳作品。此外在日益走向專業化的年代，期刊編輯必須考量評審的客觀性以及學者專家的需求，因此有些期刊開始以內聘或委外方式將稿件交由專家評審（Burnham, 1990; Rennie, 2003）。

儘管如此，同儕審查機制並未立即成為顯學，各個期刊採行同儕審查的理由分歧、起始時間不一，評審的作業方式也多有差異。Burnham（1990）以生醫期刊進行研究發現，各期刊開始實施同儕審查的年代沒有顯著的高峰期，身為期刊編輯或發行人的醫生或科學家，大都習於獨自面對專業學門上的變遷或挑戰，也很少參考或直接援引其他期刊編輯的作法。直到二次大戰後期，愈來愈多期刊編輯將稿件交由專家評審，期刊市場採用同儕審查機制的趨勢已然成形，不過同儕審查仍然沒有標準程序，而且各個期刊有不同的作業方式（Manske, 1997）。

(二) 期刊同儕審查的當代發展

當代期刊同儕審查普及化的另一股助力來自於書目計量（bibliometrics），20世紀中葉，Garfield利用引文分析提出期刊影響係數（impact factor）（Garfield & Sher, 1963），使得期刊同儕審查的效度具象化，同儕審查不再只是為期刊文獻的品質背書，也具有期刊品牌保證的功用。許多期刊資料庫亦將同儕審查列為期刊收錄的必備條件之一（OECD, 2011a），間接地促進期刊同儕審查的發展。

到了20世紀後期，書目計量的各項指標，尤其是引文分析，又以其客觀、公平及簡易的優勢，逐漸成為評鑑個人或機構研究表現的重要參考，促使學者積極在同儕審查期刊發表文獻，除了累積個人學術成就外，並可提升獲得研究獎助或擔任大學教職的機會。因此期刊同儕審查再次經由書目計量的連接，間接成為影響獎助計畫申請及大學教職聘用與升遷的重要因素（Harley & Acord, 2011; Harley et al., 2010; van Arensbergen, van der Weijden, & van den Besselaar, 2014a）。另外對於身處資訊爆炸時代的期刊讀者來說，期刊同儕審查的文獻品管功能，也是值得信賴的閱讀保證。根據英國非政府組織Sense about Science（2010）針對期刊文獻作者及評審者進行的全球性調查發現，84%的受調者認為若無期刊同儕審查機制，學術傳播將失去控制。

經過300多年的發展，期刊同儕審查已成為科學知識的守門者，受到學術社群的重

臺灣大學學術期刊資料庫

視，但是也有批評與改革的聲浪。近年來網路科技的發展使得期刊同儕審查出現許多創新作法，例如預印同儕審查的arXiv典藏庫及出版後同儕審查的PLoS電子期刊，都是相當成功的例子；後者除可讓讀者留言評論外，每篇文獻亦提供閱讀、引用及下載次數等最新統計。但是這些成功並不代表傳統期刊同儕審查即將消失，尤其在各國政府與學術界對書目計量益加重視的情況下，學者仍然必須努力在同儕審查期刊發表研究成果，以提升個人的學術成就與聲望（Harley & Acord, 2011; Rennie, 2003）。

今天期刊同儕審查所面對的是一個快速轉換的年代，電子期刊出版的發展及品質持續改進，加上人類閱讀習慣的改變，許多學者認為期刊同儕審查在不久的將來可能發生革命性的變化。Smith（2003, 2009）認為期刊同儕審查的缺點大於優點，在過去數百年之所以停滯不前，主要是缺乏競爭者，他相信未來的同儕審查作業將會更加公開與透明，而企業界的作業流程再造及持續改進等管理作法，可能全面提升同儕審查的效率、效用與公平性。

二、獎助同儕審查的起源與發展

(一) 獎助同儕審查的起源—以英美兩國為例

除了學術期刊外，研究計畫獎助機構是最廣泛採用同儕審查的單位（Bornmann, 2011b）。19世紀以來，英美等國陸續出現支持科學研究的機構，當時即經常由科學家組成的委員會協助分配獎項，例如1831年成立的英國科學促進會（British Association for the Advancement of Science），下設各種不同領域的專業委員會，負責評審各類研究申請案件。美國華府卡內基研究所（Carnegie Institution of Washington）亦於1902年設置了18個主題諮詢委員會，評審並推薦獎助申請案件；而美國國家研究委員會（National Research Council）在1919年就利用同儕審查機制決定獎學金人選（Burnham, Sauer, & Gibbs, 1987）。

基本上獎助同儕審查的發展與政府職能擴張息息相關，當各國政府開始承擔支持科學研究的責任，同儕審查就逐漸成為政府獎助經費的主要分配工具（Frodeman et al., 2012）。以英國為例，第一次大戰後期Richard Haldane（1856-1928）授命檢討政府效能，所提建議之一是將政府資助的科研計畫分為指定及非指定研究兩類；前者由政府主導獎助決策，後者則設置獨立的科學研究委員會，讓科學家共同決定研究的優先順序與經費分配，以強化科學發展的自主性，並避免來自政治人物或行政部門的干預，這就是著名的霍爾丹原則（Haldane Principle）。兩年之後（1920年）英國醫學研究委員會（Medical Research Council）成立，採用同儕審查方式分配醫學類非指定研究獎助經費（Boden et al., 1990; RCUK, 2006; RIN, 2010）。目前英國有七個不同領域的研究委員會，都是依據皇家憲章設置的準政府機構，由政府科技主管部門訂定各個委員會的經費預算與政策目標，並考核其運作績效。

臺灣大學學術期刊資料庫

根據英國醫學研究慈善協會（Association of Medical Research Charities）統計，英國各種公共基金每年發放出去的醫學研究獎助經費有20多億英磅，其中超過95%是透過同儕審查機制進行分配（Ismail et al., 2009）。

美國獎助同儕審查的發展主要在二次大戰之後，政府開始編列巨額預算獎助科學研究。與英國的發展相較，美國的獎助同儕審查在一開始就有兩種不同作法，一為法制化模式，另一為強勢計畫管理者模式（Chubin & Hackett, 1990; Guston, 2003）；前者以美國國家衛生院（National Institutes of Health, NIH）為代表，聯邦政府以法令規定NIH及其所屬機構必須採用同儕審查機制；後者主要起源於美國海軍研究署（US Office of Naval Research, ONR），該署在1940年代末期開始設立科學研究獎助基金，讓獎助計畫管理者擁有最後決定實權，而同儕審查機制則非必要選項（Burnham et al., 1987; Frodeman et al., 2012）。1950年美國國家科學基金會（National Science Foundation, NSF）成立，仿效ONR的強勢計畫管理者模式，但是也私下納入NIH的同儕審查作法（England, 1982; Frodeman et al., 2012）。

20世紀後期，美國政府強調施政績效管理，美國科技政策辦公室（Office of Science and Technology Policy）為了促進聯邦獎助機構採用同儕審查機制，自1996年起聯合管理與預算辦公室（Office of Management and Budget）共同提出年度指導綱領，要求各聯邦獎助機構增加同儕審查評選之獎助比例（General Accounting Office [GAO], 1999）。到了2000年，美國聯邦R&D預算中共有260億美元（31.4%）係以同儕審查機制進行分配（Guston, 2003）。另外美國政府獎助機構也受到較多政治團體的影響，美國國會主張的學術指定撥款（earmark）即為一例（Savage, 1999; Scarpa, 2009）；有些國會議員認為由科學家主導的獎助同儕審查存有地理偏見，要求直接依據地理區域分配獎助經費。根據美國科學促進會（American Association for the Advancement of Science, AAAS）統計，2010年美國國會R&D學術指定撥款約42.7億美元，占R&D總費用的2.8%（American Association for the Advancement of Science, 2010）。

(二) 獎助同儕審查的當前重要議題

同儕審查原本是學術界的一種封閉、自成體系的評審過程，科學家利用這個機制將學術研究品質的控管權力限縮在科學家族之內，形成所謂的科學自治。1980年代，英美等國面對龐大的預算赤字，開始檢討政府獎勵科學研究的成效與課責性，要求政府獎助機構的經費運用除了考量科學發展外，還須提出有益社會的證據，為此社會影響指標逐漸成為各國獎助同儕審查的重要評審項目之一（Cozzens, 2001; Frodeman & Briggle, 2012; Hackett, 1997; Kamenetzky, 2012; Martin, 2011; OECD, 2011a; Roberts, 2009）。

美國NSF是較早納入社會影響指標的機構之一，1997年NSF將原來的四項評審標準

臺灣大學學術期刊資料庫

簡化為智識價值（intellectual merit）與更廣泛影響（broader impact）兩項（Holbrook, 2012; Mervis, 2011）；而歐盟第5架構計畫（1998-2002）所採用的5項評審指標中，有3項與社會影響相關（Holbrook & Frodeman, 2011）。除了前兩者外，Holbrook（2010）認為美國NIH及National Oceanic and Atmospheric Administration、加拿大Natural Sciences and Engineering Research Council，以及荷蘭Technology Foundation也都有類似的做法；而最近啟動的歐盟第8架構計畫（Horizon 2020），影響力（impact）仍然是重要審查項目之一。

除了社會影響指標，創新性亦是近年來獎助同儕審查的重要議題。許多文獻指出同儕審查具有保守傾向，不利創新計畫評選（Braben, 2004, 2011; Chubin & Hackett, 1990; Gillett, 1993; Horrobin, 1990, 1996; Langfeldt, 2006; Langfeldt & Kyvik, 2011; Luukkonen, 2012; OECD, 2011a; Roy, 1985; Yalow, 1982; Ziman, 2000）。為了改進此一現象，有些獎助機構調整同儕審查的運作方式，例如要求獎助申請者詳列研究的創新性、聘請非專業人士進入評審團隊，以及強化計畫管理者的決策權限等；有些獎助機構則創設特定獎助項目，例如指定研究主題獎項、創意科學家個人獎項、跨領域或團隊科學研究獎項等，以鼓勵創新研究（Frodeman & Holbrook, 2012; Guthrie, Guérin, Wu, Ismail, & Wooding, 2013; Ismail et al., 2009; Spier, 2002b）。

然而無論是評審社會影響或創新性，都必須面對評審者選聘及評審標準訂定兩項難題（Bell, Shaw, & Boaz, 2011; van der Meulen & Rip, 2000）。學術界除了對於兩項指標的評審者選聘缺乏共識外，反對社會影響指標的學者認為，考量社會需要可能對科學自治形成干擾，而且社會影響的效用非可立竿見影，也難以推論因果（Bhattacharya, 2012; Bornmann, 2012; Bornmann & Marx, 2014; Bozeman & Boardman, 2009; Holbrook & Frodeman, 2011; Martin, 2011; Rymer, 2011; van der Meulen & Rip, 2000）；而批評創新性指標的學者則表示，所謂創新計畫的定義不夠明確，評審標準各異，而且缺乏可靠的效用考評機制（Donovan, 2011; Frodeman & Holbrook, 2012; Lal & Peña, 2013）。

三、大學教職聘用／升遷同儕審查的起源與發展

(一) 大學教職聘用／升遷的制度化—以美國為例

對於學術界人士來說，大學教職聘用／升遷同儕審查是個神祕又令人敬畏的作業，除了評審過程強調機密性外，未能獲得終身教職的老師極可能面臨失業的困境（Weiser, 2012）。相較於期刊稿件及獎助計畫評審，美國大學教職聘用／升遷同儕審查的發展較遲，主要以學術界的同儕審查為根本，並奠基於1940年以來逐漸成熟的大學教職職級（academic ranks）與終身教職（tenure）制度（Weiser, 2012）。

臺灣大學學術期刊資料庫

早期美國大學教職聘用與升遷的權力大多集中在學校行政部門，教職工作缺乏保障，甚至可能因為學校捐款人或董事會成員向校方施壓而失去工作，直到大學實施終身教職制度才改變此一情況（Brown & Kurland, 1990; Cameron, 2010; Metzger, 1990; Weiser, 2012）。1915年美國大學教授協會（American Association of University Professors, AAUP）成立，除鼓吹學術自由外，也提出教職解雇作業準則，以及倡議終身教職制度（American Association of University Professors [AAUP], 1915）。之後AAUP在1940年提出學術自由與終身教職聲明，除詳列由試用到終身教職的進程外，並強調終身教職制度是學術自由的基礎，讓教師擁有研究出版、課堂講授與討論，以及言論發表與寫作的自主空間。但是該聲明並未提及教職聘用與終身教職的評審標準，比較相關的文句是：學術自由的目標是教學與研究（AAUP, 1940; Weiser, 2012）。

經過AAUP多年的努力，目前美國大學普遍採用終身教職制度，教職工作獲得制度性的保障。不過每所學校之教職聘用／升聘同儕審查的作業方式差異甚大，甚至在同一所大學內的不同系所也可能有別，以評審者為例，有的以內部評審為主、有的強調外部評審，有的則是內外部評審兼具（Frodeman et al., 2012; Gross-Schaefer, Gala, Jaccard, & Vetter, 2015; Weiser, 2012）。1966年，AAUP提出大學院校治理聲明，呼籲由大學校務董事成員、行政管理者、教職員、學生及其他成員等共同承擔學校治理的責任（AAUP, 1966）。這股大學共同治理風潮，不但擴增了大學教職聘用／升遷同儕審查作業所涉及之利益關係者，也使得評審的決策權力結構更加複雜多元（Cummings & Finkelstein, 2012; Weiser, 2012）。

(二) 大學教職聘用／升遷同儕審查的重要議題

整體來看，大學教職聘用／升遷同儕審查的評審標準係以教學、研究與服務為主，但是自20世紀後期以來，許多國家的大學都出現偏重研究的情況（Fairweather, 2005; Greenbank, 2006; Kreber, 2002; Pratt, 1997），原因之一為各國政府高教機構評鑑大多強調書目計量指標；之二是各國政府的高教經費縮減，校務營運益形倚重校外研究獎助資源（Andersen, 2003; Becher & Trowler, 2001; Bornmann, 2013a; Greenbank, 2006; Harley et al., 2010; Waters, 2009）。在這種雙重的壓力下，學者出版期刊文獻及獲得研究獎助之能力，就成為爭取大學教職聘用與升遷的關鍵因素（van Arensbergen et al., 2014a; van Arensbergen, van der Weijden, & van den Besselaar, 2014b）。批評者認為此一趨勢形同將大學教職聘用／升遷同儕審查的工作，委由期刊出版商及研究獎助機構辦理（Boyer, 1997; Harley & Acord, 2011; Harley et al., 2010）。

除了偏重研究外，許多國家的大學教職聘用／升遷同儕審查也有忽視教學表現的情況。英國的大學教師多意識到各級政府及學校比較重視研究成績（Dearing, 1997），

臺灣大學學術期刊資料庫

並認為教學表現與教職升等之間的關聯性不大（Young, 2006）；澳洲的調查也有類似的發現，儘管大多數的學者（88.2%）支持獎勵優質教學，但是只有31.4%的教師認為教學表現有助升遷（Bexley, James, & Arkoudis, 2011）。另外Milem、Berger與Dey（2000）指出在1972至1992年間，美國四年制大學教師從事研究工作的時數顯有增加，但是在課外與學生進行諮商或互動的機會卻相對減少。

近年來，許多國家已經注意到大學校園存在重視研究而輕忽教學的情況，積極調整高等教育管理政策並鼓勵優質教學，以促進研究與教學的平衡。2010年OECD出版高等教育優質教學回顧報告，調查分析全球20國、29家高等教育機構的46項優質教學計畫方案，提供全球各大學參考。該報告並指出各國高等教育機構評鑑以及世界大學排名大多過度重視研究指標，而引發輕忽教學的批評，不過如何評鑑教學品質，也是學術界必須積極面對的議題（Hénard, 2010）。

肆、同儕審查的研究現況

同儕審查是學術界有限資源的分配機制，其操作過程卻具機密性，在公眾領域幾乎沒有任何線索，直到1980年代前後才開始有較多理性檢驗（Chubin & Hackett, 1990; Rennie, 2003; Weller, 2002）。目前同儕審查的研究以期刊稿件及獎助計畫為主，大學教職聘用與升遷的探討甚少；不過許多學者指出大多數文獻的方法論不夠嚴謹、因果推論較為薄弱，而且各篇研究結果的異質性甚大，難以通則化，建議採用後設分析法或實驗研究加以補強（Bornmann, 2011b; Bornmann, Nast, & Daniel, 2008; De Vries et al., 2009; Demicheli & Di Pietrantonj, 2007; Jefferson et al., 2007），Marsh、Jayasinghe與Bond（2011）則認為除了二手文獻的後設分析外，大規模一手評審資料的研究亦有其必要性。

一、打開同儕審查的「黑盒子」

期刊同儕審查的理性檢驗在1980年代逐漸受到重視，Rennie（2003）舉出兩項具體事證，其一為1985年*The British Medical Journal*（BMJ）總編輯Stephen Lock在英國奈菲爾基金會（Nuffield Foundation）的支持下出版第一本期刊同儕審查專書—《困難的平衡》（*A Difficult Balance*），內容分析期刊同儕審查的作業細節，並呼籲進行更多的討論與研究；其二是1989年首屆同儕審查與生醫期刊研討會在美國芝加哥舉行，之後每四年辦理一次，除了論文發表數量逐屆成長外，研究的品質與複雜度也見提升。在這股研究風潮的影響下，加上期刊市場競爭日烈，許多期刊編輯開始進行同儕審查作業的自我檢驗；而透過愈來愈多的研究，期刊編輯也益加無法專斷決定稿件的命運（Altman, 1996; Rennie, 1998a, 1998b, 2003; Smith & Rennie, 1995）。

獎助同儕審查的早期研究亦多屬自我檢驗類型，以美國為例，NIH及NSF在1970至1980年代即曾分別委託學者進行同儕審

臺灣大學學術期刊資料庫

查作業的自我檢驗（Burnham et al., 1987; Chubin & Hackett, 1990）。Chubin與Hackett（1990）將獎助同儕審查的研究分為三類並說明其特色，其一、獎助機構委託學者之研究：獎助機構指定學者並提供研究資料，因為學者類似自己人，研究結論大多偏向機構立場；其二、獎助機構贊助學者之研究：學者雖然擁有較多探索及表達的自主空間，但是研究資料仍需仰賴獎助機構提供，增加了研究的複雜性，研究結論也比較容易失真；其三、學者自發性地獨立研究：學者不但沒有獲得任何經費與社會支援，而且必須自行設法接近研究目標群體。直到20世紀末期，許多國家的公共獎助機構在政府資訊開放及政策目標管考的壓力下，為了提升同儕審查的作業品質與效能，比較願意將第一手評審檔案提供學者進行研究，有些甚至允許學者訪談評審者或觀察評審委員會的議事與決策過程（Bornmann & Daniel, 2005; Lamont, 2009; Langfeldt, 2001; Marsh, Jayasinghe, & Bond, 2008; Wood & Wessely, 2003），使得獎助同儕審查的評審作業逐步邁向開放與透明之途。

相較於期刊及獎助同儕審查，大學教職聘用／升遷同儕審查的作業程序更為複雜，除了利益關係者眾多及評審權力結構多元外（Cummings & Finkelstein, 2012），評審過程亦具機密性，相關研究一向不多（Brooks, 1988; Weiser, 2012）。但是到了1990年代前後，許多國家的教育經費縮減及學生人數減少，大學教職的穩定性受到影響，此項強調學者個人表現的審查作業開始受到重視，有關評審標準的討論也日漸增加（Becher & Trowler, 2001; Harley & Acord, 2011; Harley et al., 2010; Weinbach & Randolph, 1984），不過大學教職聘用／升遷同儕審查，迄今依然處在密閉的黑盒子之中。

二、同儕審查的研究現況

(一) 期刊同儕審查的研究現況

Chubin與Hackett（1990）將期刊同儕審查研究分為兩類，一者針對評審作業程序，探討稿件審查的準確性或評審意見的信度與效度等；一者視同儕審查為科學研究的守門者，強調評審的公平性，譴責各種類型的偏見。許多學者指出期刊同儕審查文獻的研究設計差異甚大，有針對編輯、評審者或作者進行調查；有探討評審結果與可能變項（評審者或編輯背景等）之相關性；有以編輯或評審者的審查報告進行內容分析或個案研究，不過各篇文獻大都只著重於同儕審查的某個面向，在研究方法上也各有其優點與不足（Bornmann, 2011b; Chubin & Hackett, 1990; De Vries et al., 2009; Jefferson et al., 2007）。Weller（2002）亦發現同儕審查的研究雖然橫跨各個領域，但是主要集中在醫學、社會科學及心理學，而且各學科的研究重點並不相同，例如醫學領域對於統計方法的評審方式甚為關切、心理學與社會學則強調評審者的偏見，而有關拒絕率的研究通常來自於社會科學領域。

Bornmann（2011b）的回顧文獻指出，最近十餘年的期刊同儕審查實徵研究以信度

臺灣大學學術期刊資料庫

及公平性為主，許多研究發現評審者之間的信度不高，但是評審者對於拒絕稿件的共識程度高於接受稿件；至於公平性問題雖經有些文獻證實，然而因為研究結果不一，難以通則化。另有少數文獻針對同儕審查的預期效度進行檢驗，探討是否評選出最佳稿件，有些作者發現稿件遭到拒絕之後，另在其他期刊登載的比例甚高，而且不必然是知名度較差的期刊，似乎顯示稿件審查除了基於科學品質外，也受到審查過程所在情境的影響；也有學者利用引用數據作為效度指標，比較接受稿件與拒絕稿件之影響力，大都顯現編輯的決定具有高預期效度，還有研究證實接受稿件的被引用情形也高於拒絕稿件（Bornmann, 2010）。

期刊同儕審查迄今只有少數準實驗或實驗研究，Peters與Ceci（1982）將美國12家著名期刊2至3年前出版的文獻各選1篇（共12篇），經變造作者姓名及服務機構後再投稿原出版期刊，結果只有3篇被識破，另外9篇中有8篇又再次經過同儕審查的程序，不過反而成了拒絕稿件。此篇文獻的研究設計雖然遭到學術道德的批評，但是卻被後續研究者廣泛引用，做為批評期刊同儕審查的重要佐證。另外有些期刊主動進行實驗研究，例如評審者的偵錯能力分析（Baxt, Waeckerle, Berlin, & Callaham, 1998），以及雙盲或簽名評審對評審結果的影響等，都是為了提升同儕審查的作業品質與公平性（Godlee, Gale, & Martyn, 1998; Justice et al., 1998; McNutt, Evans, Fletcher, & Fletcher, 1990; Nylenna, Riis, & Karlsson, 1994; van Rooyen, Godlee, Evans, Smith, & Black, 1998）。

近年來期刊同儕審查的發展受到網路及電子出版的衝擊，有些期刊出版者或研究機構主動嘗試新型態評審作法，例如預印、出版後或公開同儕審查等，辦理的成效不一，有的持續至今，有的半途中止（Bornmann, 2011b; Harnad, 2000; Nature, 2006; RIN, 2011）。不過許多獎助機構積極支持這些創新作法或相關研究，希望及早掌握網路時代最適當的期刊稿件評審模式（Borgman, 2007; Bornmann & Daniel, 2010; Bornmann, Marx, Schier, Thor, & Daniel, 2010; Ford, 2013; Odlyzko, 1996; van Rooyen, Delamothe, & Evans, 2010）。

(二) 獎助同儕審查的研究現況

公平性與信度是獎助同儕審查長期受到關注的議題，預期效度的研究則較少，重要研究發現分述如下：(1)公平性：許多文獻雖然已經證實存在評審者偏見，不過大多數的學者在探討公平性時，係將多階段的獎助同儕審查視為一個整體，僅由評審結果分析可能的偏見。此類研究的主要問題在於因果推論薄弱，而且各篇文獻的結論異質性甚大，難以通則化（Bornmann, 2011b; Demicheli & Di Pietrantonj, 2007）；(2)信度：獎助同儕審查評審者之間的信度過低是普遍存在的現象（Cicchetti, 1991; Goldman, 1994; Hodgson, 1997; Oxman et al., 1991），有些學者認為低信度來自於評審者的個人偏見（Eckberg, 1991; Kostoff, 1995; Opthof & Wilde, 2009;

臺灣大學學術期刊資料庫

Wessely, 1998）；有些學者卻認為評審者的看法不一致，符合科學評審的常態（Chubin & Hackett, 2003; Cole, 1992; Stricker, 1991; Wood & Wessely, 2003）；(3)預期效度：少數文獻利用獲獎及落選計畫的引用數據，探討評審的預期效度，各篇研究結果的反差甚大，沒有系統性結論；不過此類研究有其先天困難，因為落選的獎助計畫大多胎死腹中，難以追蹤其效度（Bornmann, 2011a; Langfeldt, 2006; Mutz, Bornmann, & Daniel, 2015）。

近年來獎助同儕審查研究有了大幅進展，愈來愈多政府獎助機構願意開放評審過程資訊提供研究，學者不再只是針對評審結果進行分析，並可探討各階段的審查作業細節（Bornmann, Leydesdorff, & van den Besselaar, 2010），有學者觀察獎助評審委員會的議事運作及決策過程（Lamont, 2009; Langfeldt, 2001, 2006; Luukkonen, 2012; Olbrecht & Bornmann, 2010）；有學者分析內外部評審者的審查報告及決審會議紀錄（Abdoul, Perrey, Amiel, et al., 2012）；亦有學者探討不同評審階段的評審者所使用的評審標準，以及各階段評審結果的穩定性與相關性等（Bornmann & Daniel, 2005; Bornmann, Mutz, & Daniel, 2008; van Arensbergen et al., 2014a; van den Besselaar & Leydesdorff, 2009）。目前大部分研究指出，獎助同儕審查的內、外部評審者或評審委員會成員，大多依據獎助機構所預訂的評審標準進行審查，但是每位評審者對於各項標準的看法或重視程度有別，是否因而影響評審結果，仍有待進一步討論（Abdoul, Perrey, Amiel, et al., 2012; Bornmann & Daniel, 2005; Langfeldt, 2001）。

獎助機構也進行少數實驗性研究，最著名者為美國NSF的評審者一致性實驗，研究團隊推論有一半的獲獎者是受到隨機因素的影響（Cole, Cole, & Simon, 1981）。近年來European Molecular Biology Organization進行同儕審查性別偏見檢測，得到不利於女性申請者的結論（Ledin, Bornmann, Gannon, & Wallon, 2007）；另有澳洲研究委員會（Australian Research Council）以少數專家審閱所有申請案件，發現除了評審者之間的信度提升外，在評審時間與經費上亦較為經濟（Jayasinghe, Marsh, & Bond, 2006）。

(三) 大學教職聘用／升遷同儕審查的研究現況

大學教職聘用／升遷同儕審查的研究文獻不多，主要探討評審標準、評審作業程序，以及評審決策權力結構等議題。美國當代語言協會（Modern Language Association, MLA）於2005年進行的教職升遷調查指出，美國四年制英語及外語系所在評審終身教職時，認為出版品非常重要者達75.7%，比Wilcox（1970）在40多年（1966-1967）前的調查高出兩倍強（35.4%）；該調查亦發現數位文獻或專論在評審時比較不受重視，甚至有不予承認的情況（Modern Language Association, 2006）。另有一些文獻證實研究計畫獲獎紀錄也是大學教職聘用／升遷的重要因素，例如丹麥研究委員

臺灣大學學術期刊資料庫

會Independent Research的獲獎者成為正教授的比率為16%，落選者則只有9%（Bloch, Graversen, & Pedersen, 2014）。歐盟研究委員會（European Research Council）的Starting Grants及德國研究協會（Deutsche Forschungsgemeinschaft, DFG）的Emmy Noether Programme的獲獎者，也有類似的情形（Hornbostel, Böhmer, Klingsporn, Neufeld, & von Ins, 2009; Laudel & Glaser, 2012）。

關於大學教職聘用／升遷同儕審查的評審作業程序，Weiser（2012）蒐集美國多所大學的運作方式，綜整提出五大共通點，包括評審項目（新聘或教職職級）、評審者（內部評審或外部評審）、評審者的評審範圍、評審標準，以及評審作業的機密性。他並指出內部評審一般分為系級、院級及校級，各級評審委員會的組成及標準都不相同，至於外部評審則以研究型大學較為普遍。另外有關評審決策權力結構的研究，根據德國Kassel大學International Center for Higher Education Research進行之跨國調查發現，大學教職的研究、教學與服務評鑑所涉及之利益關係者眾多，包括政府與外部利益團體、校級行政部門、系主任與院長、教師委員會或協會、同事，以及學生等六大類，而且各利益關係者對於教師評鑑結果的影響力也隨著國家與機構而有所不同（Cummings & Finkelstein, 2012）。

近年來，許多學者批評大學教職聘用／升遷同儕審查存在偏重研究的現象，並且過度重視學者之期刊文獻出版及研究計畫獲獎的能力（Andersen, 2003; Harley & Acord, 2011; Harley et al., 2010）。為了促進學術領域有關個人表現評審的公平性，歐盟第7架構支持的跨國合作研究計畫ACUMEN（Academic Careers Understood through Measurements and Norms），整合了同儕審查、書目計量及網路計量3種評審方式，提出個人學術表現評審架構—ACUMEN Portfolio，並出版詳細操作綱領，以供獎助計畫或大學教職聘用／升遷等主事機構參考運用。ACUMEN架構的第一部份是受評者經歷自述（career narrative），第二部分為專業（expertise）、產出（output），以及影響（influence）等三項次架構（ACUMEN Consortium, 2014; Tatum & Wouters, 2013）。

伍、同儕審查的展望

近半個世紀以來，除了學術界對於同儕審查研究的廣度與深度逐漸提升外，同儕審查的主事機構亦積極進行國際合作，出版評審作業的基本規範或操作指南，以強化審查的品質與公信力。在此同時，書目計量也憑藉著客觀與簡易的特質逐漸受到各界的重視，近年來更與同儕審查並列為學術評鑑的兩大主要工具。有些學者認為同儕審查與書目計量各有其優缺點，而且具有彼此互補功能，若能整合運用，應可提升學術評審作業的品質與效用。

臺灣大學學術期刊資料庫

一、同儕審查的國際合作

同儕審查的作業方式在大多數的國家均未以法律規範，而是由主事機構自行決定，或有明訂操作指南，或有依內部慣例進行，各個機構的作法甚為多元（Bornmann, 2011b; GAO, 1999; OECD, 2011a; Rennie, 2003; Weiser, 2012; Weller, 2002; Wood & Wessely, 2003）。不過多年來各國同儕審查主事機構亦積極進行國際交流，彼此分享經驗，以提升同儕審查的品質與公信力。目前期刊及獎助同儕審查都已有跨國性的合作成果，有些學者亦主張透過高等教育機構的跨國合作，共同選定大學教職聘用／升遷同儕審查的基本評審指標（Harley & Acord, 2011; Weiser, 2012）。

期刊同儕審查之跨國合作較早，國際醫學期刊編輯委員會（International Committee of Medical Journal Editors, ICMJE）於1979年出版《生物醫學期刊投稿統一規範》（*Uniform Requirements for Manuscripts Submitted to Biomedical Journals*），提供作者及期刊編輯參考。2013年該委員會大幅修訂《統一規範》並更名為《學術研究的管理、報告、編輯與出版在醫學期刊之建議規範》（*Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals*），詳列作者、評審者、編輯、出版者在同儕審查過程中的職能與責任，並強調同儕審查是科學研究過程中不可或缺的一環（International Committee of Medical Journal Editors, 2013）。委員會除呼籲生醫與健康照護期刊支持採行《建議規範》外，並定期更新內容，最新版已於2015年12月出版。

獎助同儕審查機構的跨國合作發展較遲，歐洲研究機構負責人協會（European Heads of Research Councils）及歐洲科學基金（European Science Foundation, ESF）有感於歐洲各國獎助類型與同儕審查作業的多樣性，於2010年合作推動跨國性調查，並於次年出版《歐洲同儕審查指南》（*European Peer Review Guide*）。全書的第一部分為獎助同儕審查作業概覽，說明獎助的類型與差異以及同儕審查的價值與作業方式等；第二部分則分別論述不同類型的獎助計畫，包括個人與合作研究計畫、職業生涯發展計畫、科學網絡創設與強化計畫，以及區域研究中心與研究基礎設施建置計畫等（European Science Foundation, 2011a, 2011b）。

此外美國NSF為了建立獎助同儕審查的核心價值，於2012年首度召開全球科學同儕審查高峰會，共有近50國（多為G20及OECD的會員國）的獎助機構代表與會，會中決議建立全球研究委員會（Global Research Council, GRC），並提出科學同儕審查六項原則聲明，包括專家評審（expert assessment）、透明性（transparency）、公平性（impartiality）、適切性（appropriateness）、機密性（confidentiality），以及誠信與道德考量（integrity and ethical considerations）；聲明中亦強調獎助機構作為公共資金的管理

臺灣大學學術期刊資料庫

者，必須展示它們對於送審之研究計畫具有卓越的評審能力，並確保其研究宗旨符合政策目標；而嚴謹且透明的同儕審查作業，將有助於確保政府的獎助經費應用在最能促進科學發展及解決社會問題的計畫上（Global Research Council, 2012）。

二、同儕審查與書目計量

學術界除了進行同儕審查的研究外，也積極尋求改進或替代方案。近年來期刊同儕審查受到網路科技及電子出版的影響，已出現許多創新作法而呈現多元發展。至於獎助計畫及大學教職聘用／升遷同儕審查，最常引發討論的是同儕審查與書目計量的競合關係，有些學者認為廣泛且多樣的書目計量指標有助於提升同儕審查的合理性與透明性（Bornmann, 2011a, 2013a; van Raan, 2005）。Geisler（2001）指出同儕審查與書目計量的結合應用，讓同儕審查的決定不再只是評審者的主觀意見，也加入了客觀的量化指標；但是如何將書目計量的量化資訊妥適地納入各種不同且多樣的同儕審查機制之中，是學術界面臨的重大挑戰。

(一) 書目計量與科學影響力

Alan Pritchard在1960年代提出書目計量，用統計方法呈現已紀錄的資訊，例如計算專書、文獻、出版品，以及引用的數值（Bornmann, 2013a），其中引用數據已逐漸被視為科學影響力的重要指標之一（Daniel, 2005; Garfield & Welljamsdorof, 1992; Smith, 1981）。Merton（1988）認為文獻引用在知識傳遞與擴散上具有雙重意義，在工具意義上，引用代表可能具有參考價值的資訊；而在系統意義上，引用是科學家在知識的智慧財產資料庫中，對於同儕的研究成果進行認知註記。

許多學者質疑引用的動機並指出引文分析的限制，包括引用行為偏好、引用效用可議，以及引文資料庫的正確性不足等；甚至有學者認為學術傳播系統本不完美，出版品的重要性不等同於影響力，而且被大量引用的文獻也未必一定是高品質（Bornmann, 2011a; Laloë & Mosseri, 2009; Martin & Irvin, 1983; Peters & van Raan, 1994）。不過Bornmann與Daniel（2006）回顧1960至2005年出版的引用行為實徵文獻發現，雖然大多數研究證實引用動機不只是對於科學家同儕之智識與認知影響表達認同，也受到其他非科學價值因素的影響。但是這些研究被視為缺乏信度，因為各篇的研究設計差異甚大，研究結果幾乎無法複製，而且很多文獻在方法論上有其缺陷。

Van Raan（2005）認為許多研究顯示引用動機並非如此的不同或隨機，反而在許多情況下，引文分析確實是影響力的可靠指標。支持引文分析的學者表示，儘管許多學者認為同儕審查是品質保證的象徵，但是對於大多數的人來說，同儕審查只是提供出版品給科學社群，但是接受度（引用）才是影響力的代表（Bornmann, 2013a; Shadbolt, Brody, Carr, & Harnad, 2006）。Research Evaluation and Policy Project（2005）指出雖

臺灣大學學術期刊資料庫

然研究品質的評鑑需由同儕進行，但是許多國家的政策已逐漸將影響力視為品質的代理指標。

(二) 同儕審查與書目計量的互補性

質性的同儕審查與量化的書目計量是當前學術領域的兩大評審方式，各有其支持者，也受到許多批評。許多文獻探討同儕審查與書目計量的互補關係，有些學者認為學術評鑑應始自學術價值、終於學術影響力，因此同儕審查必須結合書目計量，才得以完整呈現學術評鑑的全貌（Borgman, 2007; Bornmann, 2011a; The British Academy, 2007; Pendlebury, 2008）。Wouter（1997）以資訊流的觀點提出科學知識循環模式，分析同儕審查與書目計量的循環互動關係，認為透過書目計量之檢驗，同儕審查過程中的領域專家將難以專斷自主。

Van Raan（1996）認為書目計量指標不應該單獨使用，建議將強調研究表現的書目計量指標（尤其是引用數據）納入同儕審查的評審過程之中，以作為評審者的重要參考資訊。此一作法使得同儕審查不再只是少數評審者的意見，評審者可以透過書目計量瞭解受評者在全球研究前沿的地位、影響力與特殊性，並可進一步洞察科學傳播模式與知識散佈過程，此即所謂的「資訊充分的同儕審查」（informed peer review）。Bornmann（2013a）指出書目計量與同儕審查整合的效益有二，其一、廣泛且重要的書目計量指標，有助提升同儕審查的透明性與合理性；其二、書目計量可以檢驗同儕審查的結果，以避免學閥派系效應，而此點正是同儕審查的根本問題。他也認為書目計量指標必須經過專家評審者的挑選與解讀，才可以用來評鑑受評者的專業表現。

目前同儕審查與書目計量的整合評鑑方式，已經應用在機構層級的學術評鑑，例如歐洲許多國家採行的高等教育機構評鑑，即同時包括同儕審查與書目計量指標。但是此一方式應用在個人層級的評鑑時，就引發較多爭議，如何選擇適當的書目計量指標、如何兼顧領域的差異性，以及如何分配兩者的權重等，都是受到關切的議題，這些亦是獎助計畫與大學教職聘用／升遷同儕審查必須面對的重要議題（Abramo & D'Angelo, 2011; Bertocchi, Gambardella, Jappelli, Nappi, & Peracchi, 2015; Cabezas-Clavijo, Robinson-García, Escabias, & Jiménez-Contreras, 2013; Harley & Acord, 2011; Harley et al., 2010; Weingart, 2005）。

陸、結語

在人類的活動中，科學研究可能是受到最多檢驗與評鑑的項目之一（Laloë & Mosseri, 2009），以確保研究品質，並做為學術有限資源的分配機制。但是同儕審查有其先天限制，一者是評審乃人類的行為，容易受到人性弱點或偏見的影響；一者是評審者擁有極大的權力，但是卻強調保密性（Geisler, 2000; Wenneras & Wold, 1997; Ziman, 2000），因此招致「黑盒子」之批評，甚至有學者認為繼續採行同儕審查的唯一理由是：缺乏其他更好的方法（Kostoff, 2004; Rennie, 1986; Sieber, 2006）。

臺灣大學學術期刊資料庫

若就科學自治的精神來看，同儕審查的正當性是基於學術社群成員之間的彼此信賴與誠信。英國人文社會科學院的同儕審查報告指出，評審者在審查文件時，可以接觸到其他同儕的原始資料組、新的實徵結果，或是創新概念架構。這些資訊在商場上都被視為是商業機密，但是在學術領域中，評審者卻是扮演科學管控系統的一環，其終極目標在於快速有效地傳播研究成果，因此評審者的誠信就非常重要（The British Academy, 2007）。然而許多研究已經證實評審者存有各種偏見、評審者之間的一致性過低，以及評審效用難以證實等，必須強化同儕審查的公平性、合理性，以及透明性。

今天同儕審查已經普遍受到學術界的重視，而成為各項學術活動的主要仲裁者。不過學術領域越來越專精且複雜，研究人口也愈來愈多，在學術資源未見大幅成長的背景下，競爭將日趨激烈，同儕審查研究的益受重視將不言可喻。證諸目前許多著名同儕審查期刊的稿件接受率只有個位數字，各國政府與民間獎助機構之獲獎率也有逐年下降的趨勢（National Institutes of Health, 2013; National Science Foundation, 2011; Powell, 2010; RCUK, 2006），另外大學教職聘用與升遷的評審方式亦已受到公平性的批評。因此未來同儕審查作業勢必面對外界更多的質疑與挑戰，而建立一個持續監督、檢驗與改進的機制，應是學術界的共同努力方向（Bornmann & Daniel, 2008; Callaham, 2003; De Vries et al., 2009; Gluckman, 2012; Godlee & Jefferson, 2003; Henly & Dougherty, 2009; Hojat, Gonnella, & Caelleigh 2003; Langfeldt, 2001; Rennie, 2003）。

參考文獻 References

Abbott, A. (2008, June). *Publication and the future of knowledge*. Paper presented at the Association of American University Presses, Montreal, Canada. Retrieved from http://home.uchicago.edu/~aabbott/Papers/aaup.pdf

Abdoul, H., Perrey, C., Amiel, P., Tubach, F., Gottot, S., Durand-Zaleski, I., & Alberti, C. (2012). Peer review of grant applications: Criteria used and qualitative study of reviewer practices. *PLoS ONE, 7*(9), e46054. doi: 10.1371/journal.pone.0046054

Abdoul, H., Perrey, C., Tubach, F., Amiel, P., Durand-Zaleski, I., & Alberti, C. (2012). Non-financial conflicts of interest in academic grant evaluation: A qualitative study of multiple stakeholders in France. *PLoS ONE, 7*(4), e35247. doi: 10.1371/journal.pone.0035247

Abramo, G., & D'Angelo, C. (2011). Evaluating research: From informed peer review to bibliometrics. *Scientometrics, 87*(3), 499-514. doi: 10.1007/s11192-011-0352-7

ACUMEN Consortium. (2014). *Guidelines for good evaluation practice with the ACUMEN portfolio*. Retrieved from http://research-acumen.eu/wp-content/

臺灣大學學術期刊資料庫

uploads/D6.14-Good-Evaluation-Practices.pdf

Altman, L. K. (1996). The Ingelfinger rule, embargoes, and journal peer review—Part 1. *The Lancet, 347*(9012), 1382-1386. doi: 10.1016/S0140-6736(96)91016-8

American Association for the Advancement of Science. (2010). *Congressional R&D earmarks by agency and program*. Retrieved from http://www.aaas.org/sites/default/files/migrate/uploads/earm10c.pdf

American Association of University Professors. (1915). *AAUP's 1915 declaration of principles*. Retrieved from http://aaup.org.uiowa.edu/files/aaup.org.uiowa.edu/files/Gen_Dec_Princ.pdf

American Association of University Professors. (1940). *1940 statement of principles on academic freedom and tenure*. Retrieved from http://www.aaup.org/file/1940%20Statement.pdf

American Association of University Professors. (1966). *Statement on government of colleges and universities*. Retrieved from http://www.aaup.org/report/statement-government-colleges-and-universities

Andersen, D. L. (Ed.). (2003). *Digital scholarship in the tenure, promotion, and review process*. New York, NY: M. E. Sharpe.

Baxt, W. G., Waeckerle, J. F., Berlin, J. A., & Callaham, M. L. (1998). Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. *Annals of Emergency Medicine, 32*(3), 310-317. doi: 10.1016/S0196-0644(98)70006-X

Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed.). Buckingham, England: Open University Press.

Bell, S., Shaw, B., & Boaz, A. (2011). Real-world approaches to assessing the impact of environmental research on policy. *Research Evaluation, 20*(3), 227-237. doi: 10.3152/095820211X13118583635792

Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., & Peracchi, F. (2015). Bibliometric evaluation vs. informed peer review: Evidence from Italy. *Research Policy, 44*(2), 451-466. doi: 10.1016/j.respol.2014.08.004

Bertout, C., & Schneider, P. (2004). Editorship and peer-review at A&A. *Astronomy and Astrophysics, 420*(3), E1. doi: 10.1051/0004-6361:20040182

Bexley, E., James, R., & Arkoudis, S. (2011). *The Australian academic profession in transition*. Retrieved from http://careers.unimelb.edu.au/__data/assets/pdf_file/0003/723315/The_Academic_Profession_in_Transition_Sept2011.pdf

Bhattacharya, A. (2012). Science funding: Duel to the death. *Nature, 488*(7409), 20-22. doi: 10.1038/488020a

Biagioli, M. (2002). From book censorship to academic peer review. *Emergences: Journal for the Study of Media &*

臺灣大學學術期刊資料庫

Composite Cultures, 12(1), 11-45. doi: 10.1080/1045722022000003435

Bloch, C., Graversen, E. K., & Pedersen, H. S. (2014). Competitive research grants and their impact on career performance. *Minerva, 52*(1), 77-96. doi: 10.1007/s11024-014-9247-0

Boden, M., Ash, E., Edge, D., Reece, C., Skehel, J., & Williams, P. (1990). *Peer review: A report to the advisory board for the research councils from the working group on peer review*. Retrieved from http://webarchive.nationalarchives.gov.uk/20160217110318/http://mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003951

Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the internet*. Cambridge, MA: MIT Press.

Bornmann, L. (2010). Does the journal peer review select the "best" from the work submitted? The state of empirical research. *IETE Technical Review, 27*(2), 93-96. doi: 10.4103/0256-4602.60162

Bornmann, L. (2011a). Peer review and bibliometric: Potentials and problem. In J. C. Shin, R. K. Toutkoushian, & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 145-164). Berlin, Germany: Springer. doi: 10.1007/978-94-007-1116-7_8

Bornmann, L. (2011b). Scientific peer review. *Annual Review of Information Science and Technology, 45*(1), 197-245. doi: 10.1002/aris.2011.1440450112

Bornmann, L. (2012). Measuring the societal impact of research. *EMBO Reports, 13*(8), 673-676. doi: 10.1038/embor.2012.99

Bornmann, L. (2013a). Evaluations by peer review in science. *Springer Science Reviews, 2013*(1), 1-4. doi: 10.1007/s40362-012-0002-3

Bornmann, L. (2013b). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society for Information Science and Technology, 64*(2), 217-233. doi: 10.1002/asi.22803

Bornmann, L., & Daniel, H.-D. (2005). Criteria used by a peer review committee for selection of research fellows: A Boolean probit analysis. *International Journal of Selection and Assessment, 13*(4), 296-303. doi: 10.1111/j.1468-2389.2005.00326.x

Bornmann, L., & Daniel, H.-D. (2006). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45-80. doi: 10.1108/00220410810844150

Bornmann, L., & Daniel, H.-D. (2008). The effectiveness of the peer review process: Inter-referee agreement and predictive validity of manuscript refereeing at *AngewandteChemie*. *AngewandteChemie International Edition, 47*(38), 7173-7178. doi: 10.1002/anie.200800513

Bornmann, L., & Daniel, H.-D. (2010). Reliability of reviewers' ratings when using public peer review: A case study. *Learned Publishing, 23*(2), 124-131. doi: 10.1087/20100207

臺灣大學學術期刊資料庫

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*(3), 211-220. doi: 10.1016/j.joi.2009.10.004

Bornmann, L., & Marx, W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics, 98*(1), 487-509. doi: 10.1007/s11192-013-1161-y

Bornmann, L., Marx, W., Schier, H., Thor, A., & Daniel, H.-D. (2010). From black box to white box at open access journals: Predictive validity of manuscript reviewing and editorial decisions at *Atmospheric Chemistry and Physics*. *Research Evaluation, 19*(2), 105-118. doi: 10.3152/095820210X510089

Bornmann, L., Mutz, R., & Daniel, H.-D. (2008). Latent Markov modeling applied to grant peer review. *Journal of Informetrics, 2*(3), 217-228. doi: 10.1016/j.joi.2008.05.003

Bornmann, L., Nast, I., & Daniel, H.-D. (2008). Do editors and referees look for signs of scientific misconduct when reviewing manuscripts? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics, 77*(3), 415-432. doi: 10.1007/s11192-007-1950-2

Boyer, E. L. (1997). Scholarship - A personal journey. In C. E. Glassick, M. T. Huber, & G. I. Maeroff (Eds.), *Scholarship assessed: Evaluation of the professoriate* (Special report). San Francisco, CA: Jossey-Bass.

Bozeman, B., & Boardman, C. (2009). Broad impacts and narrow perspectives: Passing the buck on science and social impacts. *Social Epistemology, 23*(3/4), 183-198. doi: 10.1080/02691720903364019

Braben, D. W. (2004). *Pioneering research: A risk worth taking*. Hoboken, NJ: Wiley.

Braben, D. W. (2011, October). *How to identify people who might radically change the way we think about an important subject*. Paper presented at the Danish National Research Foundation Annual Meeting, Copenhagen, Denmark. Retrieved from http://dg.dk/filer/20_aars_jubilaeum/Donald_Braben.pdf

The British Academy. (2007). *Peer review: The challenges for the humanities and social sciences*. Retrieved from http://www.britac.ac.uk/policy/peer-review.cfm

Brooks, J. H. (1988). Confidentiality of tenure review and discovery of peer review materials. *Brigham Young University Law Review, 1988*(4), 706-752.

Brown, R. S., & Kurland, J. E. (1990). Academic tenure and academic freedom. *Law and Contemporary Problems, 53*(3), 325-355. doi: 10.2307/1191800

Brown, T. (2004). *Peer review and the acceptance of new scientific ideas*. London, England: Sense about Science.

臺灣大學學術期刊資料庫

Burnham, J. C. (1990). The evolution of editorial peer review. *Journal of the American Medical Association, 263*(10), 1323-1329. doi: 10.1001/jama.263.10.1323

Burnham, J. C., Sauer, J. E., & Gibbs, R. D. (1987). Peer-reviewed grants in U.S. trade association research. *Science, Technology, & Human Values, 12*(2), 42-51.

Cabezas-Clavijo, A., Robinson-García, N., Escabias, M., & Jiménez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *PLoS ONE, 8*(6), e68258. doi: 10.1371/journal.pone.0068258

Callaham, M. (2003). The evaluation and training of peer reviewers. In F. Godlee & T. Jefferson (Eds.), *Peer review in health science* (pp. 164-182). London, England: BMJ Publishing Group.

Cameron, M. (2010). Faculty tenure in academe: The evolution, benefits and implications of an important tradition. *Journal of Student Affairs at New York University, 4*, 1-11.

Chubin, D. E. (1994). Grants peer review in theory and practice. *Evaluation Review, 18*(1), 20-30. doi: 10.1177/0193841X9401800103

Chubin, D. E., & Hackett, E. J. (1990). *Peerless science: Peer review and U.S. science policy*. New York, NY: State University of New York Press.

Chubin, D. E., & Hackett, E. J. (2003, February). *Peer review for the 21st century: Applications to education research*. Paper presented at theWorkshop on the Peer Review of Education Research Grant Applications, Washington, DC.

Cicchetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences, 14*(1), 119-135. doi: 10.1017/S0140525X00065675

Cicchetti, D. V. (1997). Referees, editors, and publication practices: Improving the reliability and usefulness of the peer review system. *Science and Engineering Ethics, 3*(1), 51-62. doi: 10.1007/s11948-997-0016-4

Cole, J. R. (2000). The role of journals in the growth of scientific knowledge. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: A festschrift in honor of Eugene Garfield* (pp. 109-142). Medford, NJ: Information Today.

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago, IL: University of Chicago Press.

Cole, S. (1992). *Making science: Between nature and society*. Cambridge, MA: Harvard University Press.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science, 214*(4523), 881-886. doi: 10.1126/science.7302566

臺灣大學學術期刊資料庫

Collins, F. S., & Tabak, L. A. (2014). Policy: NIH plans to enhance reproducibility. *Nature, 505*(7485), 612-613. doi: 10.1038/505612a

Cozzens, S. E. (1999). Are new accountability rules bad for science? *Issues in Science and Technology, 15*(4). Retrieved from http://issues.org/15-4/cozzens/

Cozzens, S. E. (2001). Autonomy and accountability for 21st century science. In J. de la Mothe (Ed.), *Science, technology, and governance* (pp. 104-115). London, England: Pinter.

Cummings, W. K., & Finkelstein, M. J. (2012). Historical and comparative perspectives on the faculty role in governance. In *Scholars in the changing American academy* (pp. 111-129). Dordrecht, Netherlands: Springer. doi: 10.1007/978-94-007-2730-4_8

Daniel, H.-D. (2005). Publications as a measure of scientific advancement and of scientists' productivity. *Learned Publishing, 18*, 143-148. doi: 10.1087/0953151053584939

De Vries, D. R., Marschall, E. A., & Stein, R. A. (2009). Exploring the peer review process: What is it, does it work, and can it be improved? *Fisheries, 34*(6), 270-279. doi: 10.1577/1548-8446-34.6.270

Dearing, R. (1997). *Higher education in the learning society* [Dearing report]. Leeds, England: National Committee of Inquiry into Higher Education.

Demicheli, V., & Di Pietrantonj, C. (2007). Peer review for improving the quality of grant applications. In *Cochrane database of systematic reviews* (Issue 2). Hoboken, NJ: Wiley. doi: 10.1002/14651858.MR000003.pub2

Donovan, C. (2011). State of the art in assessing research impact: Introduction to a special issue. *Research Evaluation, 20*(3), 175-179. doi: 10.3152/095820211X13118583635918

Eckberg, D. L. (1991). When nonreliability of reviews indicates solid science. *Behavioral and Brain Sciences, 14*(1), 145-146. doi: 10.1017/S0140525X00065791

England, J. M. (1982). *A patron for pure science: The National Science Foundation's formative years, 1945-57*. Washington, DC: National Science Foundation.

European Science Foundation. (2011a). *ESF survey analysis report on peer review practices*. Retrieved from http://www.esf.org/fileadmin/Public_documents/Publications/pr_guide_survey.pdf

European Science Foundation. (2011b). *European peer review guide: Integrating policies and practices into coherent procedures*. Retrieved from https://www.vr.se/download/18.2ab49299132224ae10680001647/European+Peer+Review+Guide.pdf

Fairweather, J. (2005). Beyond the rhetoric: Trends in the relative value of teaching and research in faculty salaries. *The Journal of Higher Education, 76*(4), 401-422. doi: 10.1353/jhe.2005.0027

Fang, H. (2011). Peer review and over-competitive research funding fostering mainstream opinion to monopoly.

臺灣大學學術期刊資料庫

Scientometrics, 87, 293-301. doi: 10.1007/s11192-010-0323-4

Fletcher, R. H., & Fletcher, S. W. (2003). The effectiveness of journal peer review. In F. Godlee & T. Jefferson (Eds.), *Peer review in health science* (pp. 62-75). London, England: BMJ Publishing Group.

Ford, E. (2013). Defining and characterizing open peer review: A review of the literature. *Journal of Scholarly Publishing, 44*(4), 311-326. doi: 10.3138/jsp.44-4-001

Frodeman, R., & Briggle, A. (2012). The dedisciplining of peer review. *Minerva, 50*(1), 3-19. doi: 10.1007/s11024-012-9192-8

Frodeman, R., & Holbrook, J. (2012, March). *The promise and perils of transformative research*. Paper presented at the Workshop on the Transformative research: Ethical and societal implications, National Science Foundation, Arlington, VA. Retrieved from http://digital.library.unt.edu/ark:/67531/metadc84363/

Frodeman, R., Holbrook, J., & Mitcham, C. (2012). Part I: Defining peer review. In R. Frodeman, J. Holbrook, C. Mitcham, & H. Xiaonan (Eds.), *Peer review, research integrity, and the governance of science: Practice, theory, and current discussions*. Beijing, China: People's Publishing House.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation, 14*(3), 195-201. doi: 10.1002/asi.5090140304

Garfield, E., & Welljamsdorof, A. (1992). Citation data - Their use as quantitative indicators for science and technology evaluation and policy-making. *Current Contents, 49*, 5-13.

Geisler, E. (2000). *The metrics of science and technology*. Westport, CT: Quorum Books.

Geisler, E. (2001). The mires of research evaluation. *The Scientist, 15*(10), 39.

General Accounting Office. (1999). *Federal research: Peer review practices at federal science agencies vary* (GAO/RCED-99-99). Washington, DC: United States General Accounting Office. Retrieved from http://science.energy.gov/~/media/bes/pdf/rc99099.pdf

Gibson, M., Spong, C. Y., Simonsen, S. E., Martin, S., & Scott, J. R. (2008). Author perception of peer review. *Obstetrics & Gynecology, 112*(3), 646-652. doi: 10.1097/AOG.0b013e31818425d4

Gillett, R. (1993). Prescriptions for medical research II—Is medical research well served by peer review? *British Medical Journal, 306*(6893), 1672-1675. doi: 10.1136/bmj.306.6893.1672

Giraudeau, B., Leyrat, C., Le Gouge, A., Léger, J., & Caille, A. (2011). Peer review of grant applications: A simple method to identify proposals with discordant reviews. *PLoS ONE, 6*(11), e27557. doi: 10.1371/journal.pone.0027557

Global Research Council. (2012). *Statement of principles for scientific merit review*. Retrieved from http://www.

臺灣大學學術期刊資料庫

globalresearchcouncil.org/sites/default/files/pdfs/gs_principles-English.pdf

Gluckman, P. (2012). *Which science to fund: Time to review peer review?* Auckland, New Zealand: Office of the Prime Minister's Science Advisory Committee. Retrieved from http://www.pmcsa.org.nz/wp-content/uploads/Which-science-to-fund-time-to-review-peer-review.pdf

Godlee, F., Gale, C. R., & Martyn, C. N. (1998). Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: A randomized controlled trial. *Journal of the American Medical Association, 280*(3), 237-240. doi: 10.1001/jama.280.3.237

Godlee, F., & Jefferson, T. (2003). Introduction. In F. Godlee & T. Jefferson (Eds.), *Peer review in health sciences* (pp. xiii-xv). London, England: BMJ Publishing Group.

Goldman, R. L. (1994). The reliability of peer assessments: A meta-analysis. *Evaluation & the Health Professions, 17*, 3-21. doi: 10.1177/016327879401700101

Greenbank, P. (2006). The academic's role: The need for re-evaluation? *Teaching in Higher Education, 11*(1), 107-112. doi: 10.1080/13562510500400248

Gross-Schaefer, A., Gala, S., Jaccard, J., & Vetter, L. (2015). Being honest about tenure in the United States: The need for tenure system reform within institutions of higher education. *International Journal of Social Science Studies, 3*(4), 25-36. doi: 10.11114/ijsss.v3i4.827

Guston, D. H. (2003). The expanding role of peer review processes in the United States. In P. Shapira & S. Kuhlmann (Eds.), *Learning from science and technology policy evaluation: Experiences from the United States and Europe* (pp. 81-97). Cheltenham, England: Edward Elgar.

Guthrie, S., Guérin, B., Wu, H., Ismail, S., & Wooding, S. (2013). *Alternatives to peer review in research project funding* (RR-139-DH). Santa Monica, CA: Rand Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf

Hackett, E. J. (1997). Peer review in science and science policy. In M. S. Frankel (Ed.), *East-west dialogue on research evaluation in post-communist Europe* (pp. 51-60). Budapest, Hungary: Central European University Press.

Harley, D., & Acord, S. K. (2011). *Peer review in academic promotion and publishing: Its meaning, locus, and future*. Berkeley, CA: Center for Studies in Higher Education, UC Berkeley.

Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. J. (2010). *Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines*. Berkeley, CA: UC Berkeley, Center for Studies in Higher Education.

Harnad, S. (1996). Implementing peer review on the net: Scientific quality control in

臺灣大學學術期刊資料庫

scholarly electronic journals. In R. Peek & G. Newby (Eds.), *Scholarly publishing: The electronic frontier* (pp. 103-118). Cambridge, MA: MIT Press.

Harnad, S. (2000). *The invisible hand of peer review*. Retrieved from http://cogprints.org/1646/

Heitman, E. (2002). The roots of honor and integrity in science: Historical themes in the practical ethics of research. In R. E. Bulger, E. Heitman, & S. J. Reiser (Eds.), *The ethical dimensions of the biological and health sciences* (pp. 21-28). Cambridge, England: Cambridge University Press.

Hénard, F. (2010). *Learning our lesson: Review of quality teaching in higher education*. Paris, France: Organization for Economic Cooperation and Development.

Henly, S. J., & Dougherty, M. C. (2009). Quality of manuscript reviews in nursing research. *Nurs Outlook, 57*, 18-26. doi: 10.1016/j.outlook.2008.05.006

Hicks, D. M., & Katz, J. S. (1996). Where is science going? *Science, Technology & Human Values, 21*(4), 379-406. doi: 10.1177/016224399602100401

The Higher Education Academy. (2009). *Reward and recognition in higher education: Institutional policies and their implementation*. Retrieved from https://www.heacademy.ac.uk/sites/default/files/rewardandrecognition_2_2.pdf

Hodgson, C. (1997). How reliable is peer review? An examination of operating grant proposals simultaneously submitted to two similar peer review systems. *Journal of Clinical Epidemiology, 50*(11), 1189-1195. doi: 10.1016/S0895-4356(97)00167-4

Hojat, M., Gonnella, J. B., & Caelleigh, A. (2003). Impartial judgment by the "gatekeepers" of science: Fallibility and accountability in the peer review process. *Advances in Health Sciences Education, 8*(1), 75-96. doi: 10.1023/A:1022670432373

Holbrook, J. B. (2010). The use of societal impacts considerations in grant proposal peer review: A comparison of five models. *Technology & Innovation, 12*(3), 213-224. doi: 10.3727/194982410X12895770314078

Holbrook, J. B. (2012). Re-assessing the science - society relation: The case of the US National Science Foundation's broader impacts merit review criterion (1997-2011). Retrieved from http://digital.library.unt.edu/ark:/67531/metadc77119/

Holbrook, J. B. (2013a, October). *Peer review of team science research*. Paper presented at the Workshop on Institutional and Organizational Supports for Team Science, Washington, DC. Retrieved from http://hdl.handle.net/1853/49334

Holbrook, J. B. (2013b). What is interdisciplinary communication? Reflections on the very idea of disciplinary integration. *Synthese, 190*(11), 1865-1879. doi: 10.1007/s11229-012-0179-7

臺灣大學學術期刊資料庫

Holbrook, J. B., & Frodeman, R. (2011). Peer review and the *ex ante* assessment of societal impacts. *Research Evaluation, 20*(3), 239-246. doi: 10.3152/095820211X12941371876788

Holbrook, J. B., & Hrotic, S. (2013). Blue skies, impacts, and peer review. *Roars Transactions, a Journal on Research Policy and Evaluation, 1*(1). doi: 10.13130/2282-5398/2914

Hornbostel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & von Ins, M. (2009). Funding of young scientist and scientific excellence. *Scientometrics, 79*(1), 171-190. doi: 10.1007/s11192-009-0411-5

Horrobin, D. F. (1990). The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association, 263*(10), 1438-1441. doi: 10.1001/jama.263.10.1438

Horrobin, D. F. (1996). Peer review of grant applications: A harbinger for mediocrity in clinical research. *The Lancet, 348*, 1293-1295. doi: 10.1016/S0140-6736(96)08029-4

Huutoniemi, K., Klein, J. T., Bruun, H., & Hukkinen, J. (2010). Analyzing interdisciplinarity: Typology and indicators. *Research Policy, 39*(1), 79-88. doi: 10.1016/j.respol.2009.09.011

International Committee of Medical Journal Editors. (2013). *Recommendations for the conduct, reporting, editing and publication of scholarly work in medical journals*. Retrieved from http://www.scienceofsciencepolicy.net/sites/default/files/attachments/CAPR%20midterm.pdf

Ismail, S., Farrands, A., & Wooding, S. (2009). *Evaluating grant peer review in the health sciences*. Cambridge, England: RAND Europe.

Jayasinghe, U. W., Marsh, H. W., & Bond, N. (2006). A new reader trial approach to peer review in funding research grants: An Australian experiment. *Scientometrics, 69*(3), 591-606. doi: 10.1007/s11192-006-0171-4

Jefferson, T., Rudin, M., Brodney-Folse, S., & Davidoff, F. (2007). Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database of Methodology Reviews, 18*(2), No. MR000016. doi: 10.1002/14651858.MR000016.pub3

Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A., Rennie, D., & PEER Investigators. (1998). Does masking author identity improve peer review quality? A randomized controlled trial. *Journal of the American Medical Association, 280*(3), 240-242. doi: 10.1001/jama.280.3.240

Kamenetzky, J. R. (2012). Opportunities for impact: Statistical analysis of the National Science Foundation's broader impacts criterion. *Science and Public Policy, 40*(1), 72-84. doi: 10.1093/scipol/scs059

Kostoff, R. N. (1995). Federal research impact assessment: Axioms, approaches, applications. *Scientometrics, 34*(2), 163-206. doi: 10.1007/BF02020420

臺灣大學學術期刊資料庫

Kostoff, R. N. (2004). *Research program peer review: Purposes, principles, practices, protocols*. Arlington, VA: Office of Naval Research.

Kreber, C. (2002). Controversy and consensus on the scholarship of teaching. *Studies in Higher Education, 27*(2), 151-167. doi: 10.1080/03075070220119995

Kronick, D. A. (1990). Peer review in 18th-century scientific journalism. *Journal of the American Medical Association, 263*(10), 1321-1322. doi: 10.1001/jama.263.10.1321

Lal, B., & Peña, V. (2013, March). *Big data in evaluating transformative scientific research: Concepts and a case study*. Paper presented at the Workshop on the Big Data: Measuring the Impact of the Government's Research and Development Investments, Washington, DC.

Laloë, F., & Mosseri, R. (2009). Bibliometric evaluation of individual researchers: Not even right... not even wrong! *Europhysics News, 40*(5), 26-29. doi: 10.1051/epn/2009704

Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674054158

Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science, 31*(6), 820-841. doi: 10.1177/030631201031006002

Langfeldt, L. (2006). The policy challenges of peer review: Managing bias, conflict of interests and interdisciplinary assessments. *Research Evaluation, 15*(1), 31-41. doi: 10.3152/147154406781776039

Langfeldt, L., & Kyvik, S. (2011). Researchers as evaluators: Tasks, tensions and politics. *Higher Education, 62*(2), 199-212. doi: 10.1007/s10734-010-9382-y

Laudel, G., & Glaser, J. (2012). *The ERC's impact on the grantees' research and their careers* (EURECIA, Work package 4 summary report). Retrieved from http://www.laudel.info/wp-content/uploads/2013/12/EURECIA-WP4-report-final-Jan2012.pdf

Ledin, A., Bornmann, L., Gannon, F., & Wallon, G. (2007). A persistent problem. *EMBO Reports, 8*(11), 982-987. doi: 10.1038/sj.embor.7401109

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology, 64*(1), 2-17. doi: 10.1002/asi.22784

Levy, J. (1984). Peer review: The continual need for reassessment. *Cancer Investigation, 2*(4), 311-320. doi: 10.3109/07357908409018445

Lock, S. (1985). *A difficult balance: Editorial peer review in medicine*. London, England: Nuffield Provincial Hospitals Trust.

臺灣大學學術期刊資料庫

Luukkonen, T. (2012). Conservatism and risk-taking in peer review: Emerging ERC practices. *Research Evaluation, 21*, 48-60. doi: 10.1093/reseval/rvs001

Manske, P. T. (1997). A review of peer review. *Journal of Hand Surgery, 22A*(5), 767-771. doi: 10.1016/s0363-5023(97)80067-6

Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist, 63*(3), 160-168. doi: 10.1037/0003-066X.63.3.160

Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2011). Gender differences in peer reviews of grant applications: A substantive-methodological synergy in support of the null hypothesis mode. *Journal of Informetrics, 5*, 167-181. doi: 10.1016/j.joi.2010.10.004

Martin, B. R. (2011). The research excellence framework and the "impact agenda": Are we creating a Frankenstein monster? *Research Evaluation, 20*(3), 247-254. doi: 10.3152/095820211X13118583635693

Martin, B. R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy, 12*(2), 61-90. doi: 10.1016/0048-7333(83)90005-7

McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of peer review: A randomized trial. *Journal of the American Medical Association, 263*(10), 1371-1376. doi: 10.1001/jama.1990.03440100079012

Merton, R. K. (1942). The normative structure of science. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267-278). Chicago, IL: University of Chicago Press.

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis, 79*(4), 606-623. doi: 10.1086/354848

Mervis, J. (2011). Beyond the data. *Science, 334*(6053), 169-171. doi: 10.1126/science.334.6053.169

Metzger, W. P. (1990). The 1940 statement of principles on academic freedom and tenure. *Law and Contemporary Problems, 53*(3), 3. doi: 10.2307/1191793

Milem, J. F., Berger, J. B., & Dey, E. L. (2000). Faculty time allocation: A study of change over twenty years. *The Journal of Higher Education, 71*(4), 454-475. doi: 10.2307/2649148

Miller, D. A. (1978). Criteria for appointment, promotion, and retention of faculty in graduate social work programs. *Journal of Education for Social Work, 14*(2), 74-81. doi: 10.1080/00220612.1978.10671503

Modern Language Association. (2006). *Selected findings from the MLA's 2005 survey of tenure and promotion*. Retrieved from http://apps.mla.org/pdf/taskforcereportppt.pdf

Mutz, R., Bornmann, L., & Daniel, H.-D. (2015). Testing for fairness and predictive validity of research funding decisions: A multi-level multiple imputation for missing data approach using ex-ante and ex-post

臺灣大學學術期刊資料庫

peer evaluation data from the Austrian Science Fund. *Journal of the Association for Information Science and Technology, 66*(11), 2321-2339. doi: 10.1002/asi.23315

National Institutes of Health. (2013). *Enhancing peer review survey results report*. Retrieved from http://enhancing-peer-review.nih.gov/docs/Enhancing_Peer_Review_Report_2012.pdf

National Science Foundation. (2011). *National Science Foundation's merit review criteria: Review and revisions*. Retrieved from http://www.nsf.gov/nsb/publications/2011/meritreviewcriteria.pdf

Nature. (2006). *Overview: Nature's peer review trial*. Retrieved from http://www.nature.com/nature/peerreview/debate/nature05535.html

Nickerson, R. S. (2005). What authors want from journal reviewers and editors. *American Psychologist, 60*, 661-662. doi: 10.1037/0003-066X.60.6.661

Nylenna, M., Riis, P., & Karlsson, Y. (1994). Multiple blinded reviews of the same two manuscripts: Effects of referee characteristics and publication language. *Journal of the American Medical Association, 272*(2), 149-151. doi: 10.1001/jama.272.2.149

Odlyzko, A. M. (1996). Tragic loss or good riddance? The impending demise of traditional scholarly journals. In R. P. Peek & G. B. Newby (Eds.), *Scholarly publishing: The electronic frontier* (pp. 91-101). Cambridge, MA: MIT Press.

Olbrecht, M., & Bornmann, L. (2010). Panel peer review of grant applications: What do we know from research in social psychology on judgment and decision-making in groups? *Research Evaluation, 19*(4), 293-304. doi: 10.3152/095820210X12809191250762

Opthof, T., & Wilde, A. A. M. (2009). The Hirsch-index: A simple, new tool for the assessment of scientific output of individual scientists: The case of Dutch professors in clinical cardiology. *Netherlands Heart Journal, 17*(4), 145-154. doi: 10.1007/BF03086237

Organisation for Economic Co-operation and Development. (2011a). *Issue brief peer review*. Retrieved from http://www.oecd.org/innovation/policyplatform/48136766.pdf

Organisation for Economic Co-operation and Development. (2011b). *OECD issue brief: Research organization evaluation*. Retrieved from http://www.oecd.org/innovation/policyplatform/48136330.pdf

Oxman, A. D., Guyatt, G. H., Singer, J., Goldsmith, G. H., Hutchison, B. G., Milner, R. A., & Streiner, D. L. (1991). Agreement among reviewers of review articles. *Journal of Clinical Epidemiology, 44*, 91-98. doi: 10.1016/0895-4356(91)90205-N

Parliamentary Office of Science and Technology. (2002). Peer review. *Postnote, 182*, 1-4.

Pendlebury, D. A. (2008). *Using bibliometrics in evaluating research*. Philadelphia, PA: Thomson Scientific, Research Department.

臺灣大學學術期刊資料庫

Perper, T. (1989). The loss of innovation: Peer review in multi- and interdisciplinary research. *Issues in Integrative Studies, 7*, 21-56.

Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences, 5*(2), 187-195. doi: 10.1017/S0140525X00011183

Peters, H. P. F., & van Raan, A. F. J. (1994). On determinants of citation scores - A case study in chemical engineering. *Journal of the American Society for Information Science, 45*, 39-49. doi: 10.1002/(SICI)1097-4571(199401)45:1<39::AID-ASI5>3.0.CO;2-Q

Polanyi, M. (1962). The republic of science. *Minerva, 1*(1), 54-73. doi: 10.1007/BF01101453

Popper, K. (1961). *The logic of scientific discovery*. London, England: Routledge & Kegan Paul.

Pouris, A. (1988). Peer review in scientifically small countries. *R&D Management, 18*(4), 333-340. doi: 10.1111/j.1467-9310.1988.tb00608.x

Powell, K. (2010). Making the cut. *Nature, 467*, 383-385.

Pratt, D. (1997). Reconceptualising the evaluation of teaching in higher education. *Higher Education, 34*, 23-33.

Price, D. J. (1963). *Little science, big science*. New York, NY: Columbia University Press.

Rennie, D. (1986). Guarding the guardians: A conference on editorial peer review. *Journal of the American Medical Association, 256*(17), 2391-2392. doi: 10.1001/jama.256.17.2391

Rennie, D. (1998a). Freedom and responsibility in medical publication: Setting the balance right. *Journal of the American Medical Association, 280*(3), 300-302. doi: 10.1001/jama.280.3.300

Rennie, D. (1998b). The present state of medical journals. *The Lancet, 352*, S18-S22. doi: 10.1016/S0140-6736(98)90295-1

Rennie, D. (2003). Editorial peer review: Its development and rationale. In F. Godlee & T. Jefferson (Eds.), *Peer review in health sciences* (pp. 1-13). London, England: BMJ Publishing Group.

Research Councils UK. (2006). *Report of the Research Councils UK efficiency and effectiveness of peer review project*. Retrieved from http://www.rcuk.ac.uk/RCUK-prod/assets/documents/documents/rcukprreport.pdf

Research Councils UK. (2007). *RCUK response to the project report & consultation on the efficiency and effectiveness of peer review*. Retrieved from http://www.rcuk.ac.uk/RCUK-prod/assets/documents/documents/responsereport.pdf

Research Evaluation and Policy Project. (2005). *Quantitative indicators for research assessment - A literature review* (REPP discussion paper 05/1). Canberra, Australia: Australian National University,

臺灣大學學術期刊資料庫

Research School of Social Sciences, Research Evaluation and Policy Project.

Research Information Network. (2010). *Peer review: A guide for researchers*. Retrieved from http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/peer-review-guide-researchers

Research Information Network. (2011). *E-journals: Their use, value and impact*. Retrieved from http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/e-journals-their-use-value-and-impact

Rip, A. (2000). Higher forms of nonsense. *European Review, 8*(4), 467-485. doi: 10.1017/S1062798700005032

Roberts, M. R. (2009). Realizing societal benefit from academic research: Analysis of the National Science Foundation's broader impacts criterion. *Social Epistemology, 23*(3/4), 199-219. doi: 10.1080/02691720903364035

Roy, R. (1985). Funding science: The real defects of peer-review and an alternative to it. *Science, Technology, & Human Values, 52*, 73-81. doi: 10.1177/016224398501000309

Royal Society. (1995). *Peer review - An assessment of recent developments*. Retrieved from https://royalsociety.org/~/media/Royal_Society_Content/policy/publications/1995/10260.pdf

Rymer, L. (2011). *Measuring the impact of research: The context for metric development*. Canberra, Australia: Group of Eight Australia.

Sandström, U., & Hällsten, M. (2007). Persistent nepotism in peer-review. *Scientometrics, 74*(2), 175-189. doi: 10.1007/s11192-008-0211-3

Savage, J. D. (1999). *Funding science in America: Congress, universities, and the politics of the academic pork barrel*. New York, NY: Cambridge University Press.

Scarpa, T. (2009, May). *Assessing and advancing funding of biomedical research benchmarking: Values and practices of different countries*. Paper presented at the Sigtuna Project, Sigtuna, Sweden.

Sense about Science. (2010). *Peer review survey 2009: Full report*. Retrieved from http://www.senseaboutscience.org/data/files/Peer_Review/Peer_Review_Survey_Final_3.pdf

Shadbolt, N., Brody, T., Carr, L., & Harnad, S. (2006). The open research web: A preview of the optimal and the inevitable. In N. Jacobs (Ed.), *Open access: Key strategic, technical and economic aspects* (pp. 195-208). Oxford, England: Chandos. doi: 10.1016/B978-1-84334-203-8.50020-0

Sieber, J. E. (2006). How can we research peer review? *Nature*. doi: 10.1038/nature05006

Smith, L. C. (1981). Citation analysis. *Library Trends, 30*(1), 83-106.

Smith, R. (2003). The future of peer review. In F. Godlee & T. Jefferson (Eds.), *Peer review in health science* (pp. 329-346). London, England: BMJ Publishing Group.

臺灣大學學術期刊資料庫

Smith, R. W. (2009). In search of an optimal peer review system. *Journal of Participatory Medicine, 1*, e13.

Smith, R., & Rennie, D. (1995). And now, evidence based editing. *BMJ, 311*(7009), 826-827. doi: 10.1136/bmj.311.7009.826

Snodgrass, R. (2006). Single- versus double-blind reviewing: An analysis of the literature. *Sigmod Record, 35*, 8-21. doi: 10.1145/1168092.1168094

Spier, R. (2002a). The history of the peer-review process. *Trends in Biotechnology, 20*(8), 357-358. doi: 10.1016/S0167-7799(02)01985-6

Spier, R. (2002b). Peer review and innovation. *Science and Engineering Ethics, 8*, 99-108. doi: 10.1007/s11948-002-0035-0

Stricker, L. J. (1991). Disagreement among journal reviewers: No cause for undue alarm. *Behavioral and Brain Sciences, 14*(1), 163-164. doi: 10.1017/S0140525X00065985

Tatum, C., & Wouters, P. (2013, November). *ACUMEN Portfolio: Resources for evaluation of individual researchers*. Paper presented at euroCRIS Membership Meeting, Porto, Portugal. Retrieved from http://tatum.cc/wp-content/uploads/Tatum-and-Wouters-ACUMEN@euroCRIS-Porto14Nov2013.pdf

Travis, G. D. L., & Collins, H. M. (1991). New light on old boys: Cognitive and institutional particularism in the peer review system. *Science, Technology & Human Values, 16*(3), 322-341. doi: 10.1177/016224399101600303

Van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2014a). Academic talent selection in grant review panels. In K. Prpić, I. van der Weijden, & N. Asheulova (Eds.), *(Re)searching scientific careers* (pp. 25-54). St. Petersburg, Russia: IHST/RAS - Nestor-Historia - SSTNET/ESA.

Van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2014b). Different views on scholarly talent: What are the talents we are looking for in science? *Research Evaluation, 23*, 273-284. doi: 10.1093/reseval/rvu015

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. doi: 10.3152/095820209X475360

Van der Meulen, B., & Rip, A. (2000). Evaluation of societal quality of public sector research in the Netherlands. *Research Evaluation, 9*(1), 11-25. doi: 10.3152/147154400781777449

Van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397-420. doi: 10.1007/BF02129602

Van Raan, A. F. J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*, 133-143. doi: 10.1007/s11192-005-0008-6

臺灣大學學術期刊資料庫

Van Rooyen, S., Black, N., & Godlee, F. (1999). Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. *Journal of Clinical Epidemiology, 52*(7), 625-629. doi: 10.1016/S0895-4356(99)00047-5

Van Rooyen, S., Delamothe, T., & Evans, S. J. W. (2010). Effect on peer review of telling reviewers that their signed reviews might be posted on the web: Randomised controlled trial. *BMJ, 341*, c5729. doi: 10.1136/bmj.c5729

Van Rooyen, S., Godlee, F., Evans, S., Smith, R., & Black, N. (1998). Effect of blinding and unmasking on the quality of peer review: A randomized trial. *Journal of the American Medical Association, 280*(3), 234-237. doi: 10.1001/jama.280.3.234

Ware, M. (2013). *Peer review: An introduction and guide*. Retrieved from http://www.tandf.co.uk/journals/pdf/PRC-PeerReview-Guide-2013.pdf

Ware, M., & Monkman, M. (2008). *Peer review in scholarly journals: Perspective of the scholarly community - An international study*. Retrieved from http://publishingresearchconsortium.com/index.php/112-prc-projects/research-reports/peer-review-in-scholarly-journals-research-report/142-peer-review-in-scholarly-journals-perspective-of-the-scholarly-community-an-international-study

Waters, D. J. (2009, March 2). *Archives, edition-making, and the future of scholarly communication*. Retrieved from https://mellon.org/media/filer_public/30/9d/309de9a1-94fa-40fb-bb1f-f087333e8658/djw-archives-edition-making-2009.pdf

Weinbach, R. W., & Randolph, J. L. (1984). Ratings: Peer review for tenure and promotion in professional schools. *Improving College and University Teaching, 32*(2), 81-86. doi: 10.1080/00193089.1984.10533848

Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics, 62*(1), 117-131. doi: 10.1007/s11192-005-0007-7

Weiser, I. (2012). Peer review in the tenure and promotion process. *College Composition and Communication, 63*(4), 645-672.

Weller, A. C. (2002). *Editorial peer review: Its strengths and weaknesses*. Medford, NJ: American Society for Information Science and Technology.

Wenneras, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature, 387*(6631), 341-343. doi: 10.1038/387341a0

Wessely, S. (1998). Peer review of grant applications: What do we know? *The Lancet, 352*(9124), 301-305. doi: 10.1016/S0140-6736(97)11129-1

Whitley, R., & Gläser, J. (Eds.). (2007). *The changing governance of the sciences: The advent of research evaluation systems* (Vols. 1-26). Dordrecht, Netherlands: Springer.

臺灣大學學術期刊資料庫

Wilcox, T. W. (1970). *A comprehensive survey of undergraduate programs in English in the United States*. Retrieved from http://files.eric.ed.gov/fulltext/ED044422.pdf

Williamson, A. (2003). What will happen to peer review? *Learned Publishing, 16*(1), 15-20. doi: 10.1087/095315103320995041

Wood, F. Q., & Wessely, S. (2003). Peer review of grant applications: A systematic review. In F. Godlee & T. Jefferson (Eds.), *Peer review in health sciences* (pp. 14-44). London, England: BMJ Publishing Group.

Wouters, P. (1997). Citation cycles and peer review cycles. *Scientometrics, 38*(1), 39-55. doi: 10.1007/BF02461122

Yalow, R. S. (1982). Competency testing for reviewers and editors. *Behavioral and Brain Sciences, 5*(2), 244-245. doi: 10.1017/S0140525X00011729

Young, P. (2006). Out of balance: Lecturers' perceptions of differential status and rewards in relation to teaching and research. *Teaching in Higher Education, 11*(2), 191-202. doi: 10.1080/13562510500527727

Ziman, J. (2000). *Real science: What it is and what it means*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511541391

Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva, 9*(1), 66-100. doi: 10.1007/BF01553188

（投稿日期Received: 2015/5/19　接受日期Accepted: 2016/3/14）

臺灣大學學術期刊資料庫