Research Policy 42 (2013) 287–294

Contents lists available at SciVerse ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Does incentive provision increase the quality of peer review? An experimental study

Flaminio Squazzoni[a], Giangiacomo Bravo[b,c,*], Károly Takács[d]

[a] Department of Social Sciences, GECS Research Group, University of Brescia, Via San Faustino 74B, 25122 Brescia, Italy
[b] University of Torino, Via Sant' Ottavio 50, 10124 Torino, Italy
[c] Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri, Italy
[d] Institute of Sociology and Social Policy, Corvinus University of Budapest, Közraktár u. 4–6, 1093 Budapest, Hungary

ARTICLE INFO

Article history:
Received 28 March 2011
Received in revised form 26 April 2012
Accepted 26 April 2012
Available online 22 May 2012

Keywords:
Science policy
Peer review
Cooperation
Trust
Reputation

ABSTRACT

Although peer review is crucial for innovation and experimental discoveries in science, it is poorly understood in scientific terms. Discovering its true dynamics and exploring adjustments which improve the commitment of everyone involved could benefit scientific development for all disciplines and consequently increase innovation in the economy and the society. We have reported the results of an innovative experiment developed to model peer review. We demonstrate that offering material rewards to referees tends to decrease the quality and efficiency of the reviewing process. Our findings help to discuss the viability of different options of incentive provision, supporting the idea that journal editors and responsible of research funding agencies should be extremely careful in offering material incentives on reviewing, since these might undermine moral motives which guide referees' behavior.

1. Introduction

Although peer review is crucial for innovation and experimental discoveries in science, it is poorly understood in scientific terms. Peer review is not just important for scientists, but also for institutional agencies to allocate efficiently funds and research grants and for policy makers to guarantee that taxpayer money is well invested into a credible and well functioning system. The decisive role of peers opinion is what guarantees that scientific innovation can be experimentally pursued by scientists through a continuous, decentralized and distributed trial and error process and that science can endogenously self-regulate (although influenced by external constraints and policy guidelines) by determining scientists payoffs (Squazzoni and Takács, 2011).

With origins which dates back to 1752 when the Royal Society of London obtained responsibility for the "Philosophical Transactions", this mechanism is now under increasing strain, because of the growth of scientific publishing, the increasing complexity of research technologies and interdisciplinary collaboration in each work (Alberts et al., 2008; Grainger, 2007). Not only peer review is pivotal for scientific publications (e.g., journals and books), permitting an average of about 1,400,000 ISI journal articles published yearly (Björk et al., 2009). It is also used to allocate research funds and grants, decide about scientists recruitment and promotion and evaluate universities and research institutes productivity, when standard bibliometric criteria do not hold.

Recently, many journal editors and observers have come to the conclusion that some reform of peer review is needed and that the main problem is to increase the reliability and commitment of referees (Alberts et al., 2008; Hauser and Fehr, 2007). The problem is that, although numerous studies of sociology and economics of science have investigated certain principles and mechanisms of the reward structure in science, with important implications of peer review (e.g., Stephan, 1996), few studies have specifically investigated referee behavior and how to increase commitment. A notable exception was Engers and Gans (1998), which suggested a standard economic analytic model that looked at the interaction between editors and referees. Their aim was to understand why referees were willing to perform their task without payment and whether increasing payments to referees could improve journal quality. They showed that any improvements were so costly that they made such incentives unprofitable by generating an escalation of compensation. Indeed, although payment could potentially motivate more referees to agree to review a submission, raising the review rate meant that referees could expect to impose lower costs on the journal by refusing to review a submission. While payment raised the referees' benefit of reviewing, the effect on quality could lower the costs of declining. This implied that payment should increase

* Corresponding author at: University of Torino, Via Sant' Ottavio 50, 10124 Torino, Italy.
E-mail addresses: squazzon@eco.unibs.it (F. Squazzoni), giangiacomo.bravo@unito.it (G. Bravo), karoly.takacs@uni-corvinus.hu (K. Takács).

0048-7333/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.respol.2012.04.014

to compensate for this effect, but this reduced the need for referees to incur private costs in enhancing the quality.

On the other hand, Chang and Lai (2001) followed a similar approach to understand why certain economics journals decided to give referees some kind of rewards, such as a 1-year subscription or a discount submission fee. They concluded that, if reciprocity or reputation motives were present that influenced the relationship between journals and referees, a possible snowballing effect could emerge that increased the referee recruitment rates. If accompanied by payment, this effect could even increase the review quality.

To explore empirically this problem, we have developed an innovative experiment designed to reproduce peer review dynamics under different incentive conditions. Our findings suggest that journal editors and responsible of research funding agencies should be extremely careful in offering material incentives on reviewing, since these might undermine moral motives guiding referees' behavior. On the one hand, as there is no way for editors to dig into details about the referees' effort in due course, a problem of moral hazard by referees may arise even if material incentives are present. On the other hand, and more importantly in our view, following the motivation crowding theory, the presence of material incentives might undermine intrinsic pro-social motivations of individuals by transforming reviewing into a self-interest decision problem (e.g., Bowles, 2008; Frey and Jegen, 2001). This confirms certain arguments of the sociology and economics of science about the peculiarity of the reward structure of science and its normative foundations (e.g., Stephan, 1996) and is consistent with more recent studies on the importance of social norms for reviewing, which also emphasize the irreducible heterogeneity of norms in various scientific domains (e.g., Azar, 2008; Ellison, 2002).

The rest of the paper is organized as follows. Section 2 presents a literature review that revolves around the institutional foundations of peer review, as reflected in the sociology and economics of science literature. As we will see, although these studies are important to look at the public nature of scientific knowledge, a more focused outlook on cooperation problems at the micro level of peer review is needed to understand that reputational and material incentives might be different for the figures involved and to look at how scientists ensure the quality of the knowledge produced in this situation.

Section 3 introduces our idea that the quality of peer review depends on a cooperation problem between editors, authors and referees where conflicting interests, cheating and moral hazard are all possible. Following game-theory literature on cooperation in experimental behavioral sciences (e.g., Gintis et al., 2005; Gintis, 2009), we have focused on trust, incentives and social norms. We have proposed a modified version of the investment game—i.e., a standard experimental framework (see Berg et al., 1995)—which looks at the triadic interaction between editors, authors and referees and allows us to test various incentive schemes. More specifically, our aim is to test whether material incentive provision can increase cooperation between everyone involved in peer review. While existing literature on peer review mostly takes an empirical, case-based approach (e.g., Bornmann, 2011), our idea is to look at the essential mechanisms of peer review through an abstract model that can be tested in the laboratory. This also makes a difference with the few existing economic studies on referee behavior mentioned above, which did not consider realistic and testable behavioral foundations (Chang and Lai, 2001; Engers and Gans, 1998). Moreover, this approach allows us to disentangle peer review mechanisms and to verify the impact of various interaction conditions. This also allows us to evaluate certain measures frequently recurring in the debate on peer review reform among editors of top journals (e.g., Alberts et al., 2008).

Finally, Section 4 illustrates the results of our "peer-review game" while Section 5 discusses them.

2. Science institutions and peer review

The idea that scientific knowledge is a public good and that scientists developed a normative system particularly suitable for its production, which is different from typical market and technology incentives, was suggested by Merton (1942, 1957), Nelson (1959) and Arrow (1962). In general, these classical studies argued that competitive markets provide poor incentives for scientific knowledge production as providers cannot appropriate the benefits derived from use. Moreover, being puzzle solving and discovery so intrinsically rewarding for scientists, the behavior of scientists cannot be understood as a typical maximization problem as the price of the good "knowledge" strongly depends on the preferences of the producer (Pollak and Watcher, 1975).

Dasgupta and David (1994) David (2004) followed this starting point and suggested an institutional perspective by arguing that science and technology should be seen as alternative knowledge production systems based on distinctive social institutions, i.e., distinct values, social norms and rewards. In their view, the "Realm of technology" was inhabited by secrecy, privatization and protection of knowledge, which ensured that knowledge could intercept market rewards. On the other hand, the values of openness, communitarism, disinterestedness and universalism were functional to the development of the so-called "Republic of science", where competition was based on priority and rewards followed reputational credit accumulation in the public sphere.

In an influential review of the economics literature on science, Stephan (1996) argued that these institutional features explain why a reward structure based on "non-market-based incentives" evolved in science that encouraged the production of the public good "knowledge". Her argument was that as scientists compete for priority in a context of possible mutual discoveries, they are pushed to share knowledge in a timely fashion. This generates positive externalities, such as the appropriability of knowledge by others and its growing value through multiple uses, which give rise to reputational credits for knowledge providers, such as publications and citations, which in turn fuel new knowledge production, e.g., access to new research funds for highly reputed scientists. Therefore, in this view, the fact that scientists' careers depend on reputational credits, which are built upon publications and citations, explains the public nature of knowledge and ensures a solution of the appropriability dilemma inherent in the creation of any public good.

Furthermore, these studies suggested that science can avoid the classical tragedy of the commons thanks to the strength of the intrinsic motivation of scientists (e.g., Dasgupta and David, 1994; Stern, 2004). This seems to be even corroborated by recent empirical findings. Using data survey from over 400 science and engineering PhD students in North Carolina, a recent study emphasized that students who opted for an academic career differed from those who followed a career in the private sector. Indeed, the former showed a stronger "taste for science" and a weaker concern for salary and access to resources than the latter (Roach and Sauermann, 2010; see also Lacetera, 2009). This confirmed certain results of an influential empirical survey on multiple job offers to post-doctoral biologists in the US, where it was found that wages and science were negatively correlated, so that scientists seemed to "pay" to become scientists (Stern, 2004). Similar differences of motivations and attitudes between academic scientists and private researchers were recently found also by Häussler (2011), who built a dataset of 1353 academic and 341 industry-based bio-scientists. Her results showed that academic scientists conformed to the norm of open science and shared their knowledge even when sharing