Counterintuitive effects of incentives?

Peter van den Besselaar,1,* and Ulf Sandström2

1Department of Organization Sciences & Network Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands and 2Department of INDEK, KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden

*Corresponding author. Email: p.a.a.vanden.besselaar@vu.nl.

Abstract
A recent paper in this journal compares the Norwegian model of using publications counts for university funding with a similar intervention in Australia in the mid-1990s. The authors argue that the Norwegian model (taking into account the quality of publications) performs better than the Australian (which did neglect paper quality other than being peer reviewed). We argue that these conclusions are in contrast to the evidence provided in the article, and therefore should be considered incorrect.

A recent paper in this journal argues:

The experiences from Australia where undifferentiated publication counts were linked to funding of universities in 1993 is well known. Publication activity increased, but the largest increase was in lower-impact journals, leading to a general drop in overall citation impact for Australia. The experience from Australia has been a warning for what would most likely happen if funding were linked to publication activity. Nevertheless, in 2005, a performance-based model based on differentiated publication counts was implemented in Norway. The model was specifically developed to counter adverse effects like those identified in the Australian case. (Schneider, Aagaard and Bloch 2016)

Basically, the Norwegian system is counting research output in ‘quality classes’ for two quality levels (Table 1) and four publication types (Table 2). The higher Level 2 includes approximately 8% of the journals and about 5.5% of the publishers of books and proceedings. The criteria for the levels and the weights of the publication types seem arbitrary. The Norwegian model does take into account the number of authors per publication. If three authors publish an article in a Level 2 journal, each author receives 1 point. If the same authors publish in a Level 1 journal, they receive each 0.33 points.

Schneider and his colleagues claim that the system indeed resulted into a growing productivity, without negative effects on quality (here defined as impact); both the Norwegian system and the Australian system resulted in higher output, but in Australia the citation impact went down, whereas in Norway it remained stable—so the Norwegian system does function better. As the ‘Norwegian model’ is currently under consideration or being implemented in several countries (such as Denmark; Finland; Sweden; Belgium), it is—according to the authors—an important finding as it would show that the Norwegian model is an effective science policy instrument, as it realizes the goals without negative side effects.

However, the evidence presented in the article does not support the authors’ conclusions. The authors have several graphs about Australia in their article, and the core of the evidence is in Fig. 7 (reproduced here as Fig. 1). This figure reveals no declining impact in the Australian case, as immediately after the output-based funding was implemented in 1993,1 the decline of Australian performance stopped and the impact of the Australian system started to increase again.2 This holds especially for the top 10% highly cited papers (Squares), and the less meaningful average impact indicator starts to increase shortly later (Triangles). Following these findings of Schneider et al., we would argue that the article shows that the Norwegian model did not reach its goal (‘specifically developed to counter adverse effects like those identified in the Australian case’): the Australian system not only had a positive effect on output level,

Table 1. Classification of media in the Norwegian Model in 2015

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journals</td>
<td>22,543</td>
<td>2,042</td>
</tr>
<tr>
<td>Publishers</td>
<td>1,484</td>
<td>86</td>
</tr>
</tbody>
</table>

Table 2. The value of publications in the Norwegian Model in 2015

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal article</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Chapter in book</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>Proceedings</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>Book</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>
but in contrast to the Norwegian model, it also had a positive effect on impact. Even stronger, both the Australian and the Norwegian policy interventions were followed by a trend reversal—but in different directions. In Australia, the downward impact trend was replaced by an upward trend; whereas after the Norwegian policy intervention, the upward development of impact seems to end.

Why did the authors observe a decline in Australian impact, whereas the presented data (reprinted as Fig. 1) in our view do not support this? The original findings about the Australian funding policy introduced in the early 1990s did fit very well with intuition and expectations: if you ask for more output without demanding quality, you get more output without quality! These conclusions were quickly taken up in the science policy research community, and Australia was often a more specialized audience. In the Norwegian case, we saw that the distribution of papers over the impact classes did not change. In the Australian case, the increased output was published in on average lower impact journals, but it had overall a larger impact, which supports our argument.

We conclude that the evidence in Schneider, Aagaard and Bloch (2016) shows that the Norwegian system stimulating output growth while taking ‘quality classes’ of this output into account performs less than the Australian system focusing only on output. Science policy makers may rethink what this means.

Notes
1. There is some discussion about when the policy was implemented (Van den Besselaar, Heyman and Sandstrom 2016, 2017), but we follow on this point here the same date as Schneider, Aagaard and Bloch 2016. But even if one would date the new policy 2 years earlier, the figure reveals that impact decline in the 1980s was stopped, and at the at the very least shortly after 1991 Australian impact started to rise.
2. These findings do confirm our own findings about the Australian case we reported elsewhere (Van den Besselaar, Heyman and Sandstrom 2016, 2017).
3. Or in case of books: the standing of the publisher.

References


