‘Dirty work’, but someone has to do it: Howard P. Robertson and the refereeing practices of Physical Review in the 1930s

NOTES AND RECORDS

Notes Rec. (2016) **70**, 151–174
doi:10.1098/rsnr.2015.0022
Published online 27 January 2016

'DIRTY WORK', BUT SOMEONE HAS TO DO IT: HOWARD P. ROBERTSON AND THE REFEREEING PRACTICES OF *PHYSICAL REVIEW* IN THE 1930S

by

ROBERTO LALLI*

Max Planck Institute for the History of Science, Boltzmannstrasse 22, D-14195 Berlin, Germany

In the 1930s the mathematical physicist Howard P. Robertson was the main referee of the journal *Physical Review* for papers concerning general relativity and related subjects. The rich correspondence between Robertson and the editors of the journal enables a historical investigation of the refereeing process of *Physical Review* at the time that it was becoming one of the most influential physics periodicals in the world. By focusing on this case study, the paper investigates two complementary aspects of the evolution of the refereeing process: first, the historical evolution of the refereeing practices in connection with broader contextual changes, and second, the attempts to define the activity of the referee, including the epistemic virtues required and the journal's functions according to the participants' categories. By exploring the tension between Robertson's idealized picture about how the referee should behave and the desire to promote his intellectual agenda, I show that the evaluation criteria that Robertson employed were contextually dependent and I argue that, in the 1930s, through his reports the referee had an enormous power in defining what direction future research should take.

Keywords: editorial policies; peer-review system; general relativity; *Physical Review*; Howard P. Robertson; John T. Tate

INTRODUCTION

Many criticisms have been levelled at the efficacy of the peer-review system, and scholarly debates have ended without reaching any agreement about how it should be implemented in practice.[1] Nonetheless, most scientists credit the editorial peer review as an indispensable element of good scientific practice. In public debates about controversial statements, the role of editorial peer review is so central that this practice might be considered one of the few elements that the methodologies of different scientific disciplines have in common, in natural, social and human sciences.[2] In the past 40 years an increasing number of historians, sociologists and philosophers of science have emphasized that there is no such thing as the unity of sciences, not even at a purely methodological level.[3] In a context in

*rlalli@mpiwg-berlin.mpg.de

which methodological plurality becomes increasingly visible, the peer-review system often assumes the role of the unifying *modus operandi* of the entire realm of scholarship, and sometimes it is even used as a definition of scientific method *per se*.[4]

Whether scientists really believe that the practice of peer review is what makes a particular endeavour scientific, or whether the continuous reference to the relevance of editorial peer review to validate scientific results is more of a rhetorical device employed in public debates to persuade their audience that there is a clearly defined boundary between science and non-science, it is certain that as a practice, peer review has a fundamental role in the daily activities of virtually every practitioner of science.[5] It is well accepted that some sort of peer-review process is an integral part of what might be considered the 'standard model of the scientific periodical'.[6] Scientific journals must label the work they publish as having been 'peer reviewed' if they are to be considered as legitimate venues for the publication of certified scientific results. The practice is so widespread that even journals devoted to publishing research endeavours at the very fringe of the scientific terrain proudly declare that their published papers have been peer reviewed.[7]

Studies on the contemporary peer-review system have underlined that at present the practice is not as well defined as it may appear. What is called peer review is actually a patchwork of different practices, with different aims, procedures and performance records.[8] If one tries to identify the elements that are part of what the sociologist of science Joanne Gaudet has recently christened 'the *traditional* peer review',[9] one can see that the paradigmatic aspects of today's peer review are all an outcome of historical evolution and are not as widespread as commonly perceived.

Even the term 'peer review' itself is the result of a historical process. In fact, not until the late 1960s was the word 'peer' introduced to define the refereeing system.[10] The word seems to provide a normative stance to the practice by indicating that the submitted research endeavour is to be evaluated by experts in the field, who should be, at least in principle, on the same social level in the 'stratification system' of a specific scientific community in the sense employed by Harriet Zuckerman and Robert Merton in their path-breaking sociological study of the editorial patterns of evaluation in *Physical Review* conducted in the 1960s. In that study, the authors defined four different models that depend on the social relationship between authors and referees in the stratification system of the physics community: the 'oligarchical model', the 'populist model', the 'egalitarian model' and the 'model of expertise'.[11] Within Zuckerman and Merton's interpretative scheme, the name peer review would seem to imply that the pattern of evaluation should be inherently 'egalitarian', but it is well understood that this idealized view is not easily implemented in practice.

Despite its central position in contemporary scientific practice, the historical evolution of peer reviewing has so far received little attention from professional historians of science. Apart from illuminating studies on the genesis of institutionalized refereeing systems in royal academies at the end of the seventeenth century and pioneering attempts to delineate the editorial evolution of the peer-review practice, only recently have historians of science begun emphasizing the need to build a more complete picture of the historical evolution of this all-encompassing practice.[12] Taken together, the main results of these investigations offer the general view that the historical evolution of the practice has been fragmented, both spatially and temporally, dependent as it was on local contexts, editorial choices and disciplinary traditions. There was no straightforward evolution of this practice from the first historical cases of external refereeing in *Philosophical Transactions of the Royal Society* to modern peer reviewing. As a consequence, in building a big picture of

the history of refereeing practices it seems necessary to focus on significant case studies that shed light on key aspects of the way in which these practices evolved.[13]

One of these significant case studies is the evolution of the refereeing practice within the journal *Physical Review* in the 1930s. As is well known, the publishing venues of the American Physical Society (APS)—*Physical Review* and its various sister journals and sub-journals—have held a dominant position in the cluster of physics publishing venues since before World War II, well before external refereeing became systematic as an indispensable process to validate knowledge claims.[14] As shown in a recent historical study by David Kaiser, *Physical Review* began employing external refereeing in a systematic manner for all the submitted manuscripts only in 1960.[15] Zuckerman and Merton's sociological analysis and Kaiser's historical investigations provide a fully fledged picture of the evolution of the refereeing practices and the establishment of formal peer review within *Physical Review* in the postwar period. The prewar period, in contrast, remains largely unexplored, leaving historians without a clear understanding of how the practice evolved in the crucial period during which the journal was undergoing major transformations by becoming one of the most important physics periodicals in the world.

How did the refereeing practices within *Physical Review* change during this dynamic period? How was this transformation connected to the journal's attempts to strengthen its international reputation? How was the role of the referee understood and embodied by the editors and the referees? What did the parties involved think the journal's function should be? Which were the criteria followed by the editors and referees alike for assessing the validity of a knowledge product? To what extent did local traditions and networks influence the fate of a manuscript? Which were the main epistemic virtues that a referee should possess according to the participants?

The endeavour to answer these questions is particularly challenging because no editorial archival repository contains the correspondence between the journal's editors and the referees before 1938. The correspondence about submitted manuscripts is scattered in referees' personal papers, and only a small number of these letters seem to be available for historical scrutiny. To my knowledge, the most complete collection of referee reports and related correspondence concerning the period before World War II is stored in the archived correspondence between the editors of *Physical Review* and the American mathematical physicist Howard P. Robertson, on which I base the present study.

Robertson was the main referee of *Physical Review* for papers concerning the general theory of relativity and related subjects from 1930 through the 1940s. The complex mathematical structure of general relativity made it difficult for most physicists to evaluate the submitted papers in that field, and Robertson came to be considered the major American authority for evaluating such manuscripts. The field of general relativity, moreover, was highly unstable at that time and its disciplinary boundaries were blurred between physics and mathematics. In their correspondence, the editors of *Physical Review* and Robertson had then to evaluate carefully the relevance of submitted manuscripts for the field of physics. The entire correspondence between the editors, Robertson and, in some cases, the authors has survived, revealing many features of the steps undertaken to implement the refereeing in the period under consideration. Moreover, Robertson is an interesting case because he had to be very explicit about the criteria according to which the decision about the publication of manuscripts should be taken.

Although his field of expertise made Robertson a special case not suitable for unproblematic generalization, his long-lasting and continuous correspondence permits a

broad understanding of the refereeing practices, of the journal's function and of the power structures embedded in this activity and also of their transformation in the 1930s. A comparison with a few other referee reports confirms that the dynamics of the triangular relationship author–editor–referee emerging in the correspondence between Robertson and the editors of *Physical Review* followed a quite general pattern.

To make the different aspects of the historical transformation of the refereeing practice in *Physical Review* and its broader implications emerge, this paper is structured as follows. The first section presents an overview of Robertson's scientific activities in connection with the status of relativity theory at that time. The second section focuses on the transformations of the American physics community and more specifically on the structural changes in the publication venues of the APS. In the third section the refereeing practices of *Physical Review* and their historical development are explored in connection with broader contextual changes. The fourth section concerns the relationship of the scientist with the activity of refereeing in a period when the practice was not accepted by every author as a step between the writing of a paper and its final publication in the society's journal. This investigation, in turn, sheds light on the self-perception of the referee as a *scientific persona*, with the focus on the epistemic virtues he should possess, as well as on the function of the journal according to the participants' categories.[16] By exploring the tension between Robertson's idealized picture about how the referee should behave and the desire to promote his intellectual agenda, I show that the evaluation criteria that Robertson employed were contextually dependent. In the conclusion, I argue that in the 1930s the few referees of *Physical Review* had enormous power in promoting their own agendas through the activity of refereeing.

Howard P. Robertson: an authority in a marginal field

In the 1930s general relativity was in the middle of what has been called its 'low-water-mark' period.[17] Only a few scientists worked on what many physicists considered a mathematical theory with little connection, if any, to experimental activities. It is possible to classify the branches of theoretical physics research related to general relativity in three different categories: (i) the extension of the theory towards the formulation of a unified field theory; (ii) the application of general relativity to a theory of the Universe, that is, cosmology; and (iii) more marginally, attempts at quantizing Einstein's gravitational equation.[18] Robertson made important contributions only to relativistic cosmology, but he acted as a referee for all the above-mentioned research fields. Since 1930 his advice was regularly requested to judge the submitted manuscripts in these research fields, as if Robertson were the main authority within the APS. The reason why Robertson came to hold this position was probably his strong mathematical training in both American and European universities—after which he had become one of the most authoritative American experts in differential geometry and group theory—as well as his strong interest in the foundations of physics and in the application of abstract mathematics to physical problems, especially in the field of general relativity, features that together made him a unique figure within the membership of the APS at that time.[19]

At first he was an external referee, and then from 1937 to 1940 he served as member of the editorial board.[20] The kinds of paper that were sent to him for refereeing did not change after he became a member of the editorial board, but he received them more frequently. After

1940 he continued to receive various requests to review submitted manuscripts, to which he often agreed.

In the early 1930s his connection with the APS publication activities was increased by his commitment to write a review on relativistic cosmology for *Reviews of Modern Physics*, which would later become important as the major reference on the state of the art in that field. The Irish mathematician John L. Synge recognized the relevance of Robertson's review in providing a standard reference for further developments by stating that Robertson 'had done something very useful in coordinating the entire field of cosmology'.[21]

Robertson was in fact deeply involved in the effort to define epistemological and methodological norms in the field of physical cosmology. Given its peculiar status within the natural sciences and its blurred disciplinary boundaries, most of the physics community perceived physical cosmology as a marginal research field in comparison with other branches that had a stronger relation with experimentation and observation, such as quantum mechanics and nuclear physics.[22] Within these efforts, Robertson took special care in his refereeing activity, which became part of his attempts at setting the standards of his field of expertise and at strengthening the position of his field within the American physics community.

Structural changes in *Physical Review*

In 1913 the APS had taken over *Physical Review* with the ambition that the society's management would help to strengthen the reputation of the journal.[23] The attempt to increase the prestige of the journal remained initially confined to the national context. In German academic settings, the journal was still considered a minor publication as late as 1927. Less than 10 years later, the international standing of the journal was radically increased: in the mid 1930s *Physical Review* was already regarded as one of the leading physics publications in the world.[24]

Between the end of the 1920s and the beginning of the 1930s, the APS underwent far-reaching changes, which in part reflected broader socio-political–economic transformations both in the USA and in Europe. During this period its membership grew steadily and it gained prominence within the international physics landscape in view of the increasing relevance of both experimental and theoretical activities of physicists working in the USA.[25] The growth of the society, the increasing specialization of physics practitioners and an unprecedented separation between theoretical and experimental physics were all posing enormous challenges for the editorial board of the APS's major publication venue.

After the APS assumed responsibility for the publication of *Physical Review*, the journal was managed by an editorial board consisting of one managing editor—who received a salary from the society—and nine other members, all elected by the society members for a three-year term. The historical sources do not clarify the exact function of the editorial board, which was initially to provide 'assistance and advice' to the editors, but it seems evident that the managing editor held nearly uncontested power in all the major decisions concerning the administration of the journal and its editorial policies.[26] Although the members of the editorial board were not eligible for a second term, the managing editor could, and usually did, maintain his position for many years. The managing editor had then the possibility of realizing his vision of the journal, thus providing strong editorial continuity during his period of management.

John T. Tate, who was the managing editor of *Physical Review* from 1926 to 1950, devoted himself completely to the success of the journal, and the editorship became his central scientific activity after his election.[27] As a response to the changing social composition of the community of which the journal was the main publishing venue and to its increasing international reputation, starting from the late 1920s, Tate put in place a general strategy of diversification and multiplication of the editorial production of the APS. Supported by the Council of the APS, Tate was able to promote radical changes in a relatively short time. In 1929 Tate advocated the establishment of what was later called *Reviews of Modern Physics*—a new journal entirely dedicated to the publication of reviews of recent work in various fields of physics, which would become a dominant publication venue in physics from the 1930s onwards.[28] Recognizing that speed of publication was becoming a pressing requirement for the increasing American physics community, which was in fierce competition with European physicists, in July 1929 Tate began publishing a new section of *Physical Review* entitled 'Letters to the Editors', probably inspired by the success of correspondence columns in British scientific journals such as *Nature* and *Philosophical Magazine*.[29] When Tate officially announced the establishment of this section, he stated that accepted communications would be published within two weeks from the submission date, although this promise was not always fulfilled. A third significant editorial change occurred in 1931, when a new monthly publication sponsored by the APS was launched: the journal *Physics*, which was devoted to applied physics and would become *Journal of Applied Physics* in 1937. The growth of editorial work and the strategy of diversification also led to some relevant changes in the handling of the submitted papers. Apparently, one person could no longer be sufficient to carry the administrative work of the APS journals alone. In 1931 J. William Buchta, a colleague of Tate's at the University of Minnesota, became assistant editor and continued to work as editor of APS periodicals until 1957.[30]

The implementation and transformation of the refereeing practices during the 1930s have to be understood within this context. Tate, and later Buchta, had to face the rapid increase in the editorial activities within the APS. At the same time, economic constraints worsened by the Great Depression imposed serious limitations on the space available for the publication of the papers within the pages of *Physical Review*.[31] On the other hand, Tate had to, and wanted to, establish procedures for supporting the rapid growth of the journal's prestige internationally and for facing the increasing specialization of the authors, which seemed to require an external expertise for the evaluation of manuscripts in fields in which Tate could not provide an informed judgement.

Attempts at standardizing the refereeing procedure in the 1930s

A set of tables presumably drafted by Buchta to record the editorial path of the submitted manuscripts from 1932 to 1937 shows that during these years little more than half of the submitted manuscripts were sent out for refereeing.[32] Although it is not clear when Tate began employing external refereeing, the documents show that there was a substantial employment of the practice in the 1930s, greater than in the immediate period after World War II.[33] The referees were mostly members of the editorial board, which implies that the refereeing was one of their tasks, and perhaps the major one. However, other members of the APS with a recognized expertise in a particular field also appeared in the list of

referees, although more infrequently. Furthermore, physicists who had once been members of the editorial board continued to serve as referees, even after they had left their official position at the journal.

From 1931 onwards, the establishment of the American Institute of Physics—incorporating five societies (the APS, the Optical Society of America, the Acoustical Society of America, the Society of Rheology and the American Association of Physics Teachers), each with its own publication venues—enlarged the network of editors who could be contacted to decide the more suitable journal for a specific piece of written knowledge. In the 1930s, to summarize, the network of referees and physicists able to assess the validity of a manuscript and advise on its editorial destination was not large. Nonetheless, it was slowly increasing, growing beyond the small circle of editorial board members.

The investigated correspondence and the tables allow the main steps of the refereeing process to be clarified from the manuscript's submission up to its final publication or, more rarely, rejection. Until the early 1930s, Tate tended to read all the manuscripts to evaluate whether to publish the paper or submit it for further refereeing. In the latter case, Tate asked for comments and suggestions from only one referee, who was requested to send the report within a certain timespan. (This timespan was 10 days in 1937.)[34] Only rarely did Tate later require a second review, which happened when a second opinion was directly requested or suggested by the first referee. If we assume that Tate was the only one who judged the manuscripts when the referee was not indicated in the tables, as seems plausible, then we can deduce that Tate never rejected a paper without asking for a referee report, at least in the period covered by the conserved tables. On the one hand this indicates that the journal had a very low rejection rate, which according to Zuckerman and Merton might imply that the editors were working under the assumption that a paper is publishable unless proved otherwise.[35] On the other hand this also suggests that the refereeing process could have been employed by Tate to give the anonymous authorities the responsibility to reject a paper so as to keep the role of managing editor out of possible controversies. In fact, the model of single-blind review was almost always followed, except when the referees voluntarily disclosed their identity by keeping in touch with the authors to discuss interesting matters.[36]

Robertson's entry into this activity seems to confirm that the transfer of responsibility from the editor to the referee in controversial cases was, at least in the early 1930s, one of the motivations behind the employment of external refereeing. In 1930 Tate first included Robertson in his network of referees when he asked him to help evaluate a couple of letters to the editor by the English physicist William Band, putting forward a novel approach to the unified field theories. Given the 'confused state of the problem' in unified field theories, as Robertson stressed, an evaluation of the letters was very difficult because it required a delicate balance between very different criteria.[37] In Robertson's view, the search for a unified field theory was in such a state that one could hardly say that a particular investigation was wrong in a straightforward manner. Robertson had to express clearly the motivations as to why one of the letters was worthy of publication whereas the second one, he claimed, was 'valueless'.[38] Moreover, his arguments should have permitted Tate to find a way to convince the author of the validity of the referee's judgement. After a series of exchanges, Tate heartily thanked Robertson for the 'dirty work' he had done in the case of Band's letters to the editor.[39] The use of the idiomatic expression 'dirty work' to describe what might seem a simple activity of refereeing implies that Tate was thanking Robertson for having assumed the unpleasant job of

criticizing a manuscript so that Tate could shift the responsibility for its rejection to an anonymous referee. Tate's use of this expression, along with the fact that Tate apparently never rejected a paper without having first sent it to a referee, supports the claim that Tate was also using the refereeing to maintain the role of the managing editor 'clean' as far as the decision to reject manuscripts written by his peers was concerned.

The relationship with the authors of the submitted manuscript was in fact a motive of concern for the editors, and the referee had a major role in shaping this relationship. If the paper was sent out for refereeing, the editor's response to the author strongly depended on the referee's assessments. Once Tate had requested help from a referee, he tended to follow his suggestions with some degree of arbitration. However, the practice itself was rather complex. Because the referee only rarely sent the editor a separate impersonal document to forward to the author, the editor had to write a different letter himself, employing some of the referee's arguments.[40] Sometimes, the referee explicitly indicated which remarks should be included in the editor's response to the author.[41] More often, the editor had both the freedom and the duty to choose from the referee report what he considered appropriate. This triangular process made Robertson free to use ironic and dismissive expressions in assessing the value of a research endeavour. This process, in turn, could also create many problems for the editors, who had to do very careful work to modify dismissive expressions included in sharp referee reports.

Sometimes errors occurred, which could strongly affect the relationship between the journal and the physicists who were willing to have their name printed in its pages. Sending the response by the theoretical physicist Ludwik Silberstein to Robertson's negative assessment of his manuscript, Tate explained that 'in paraphrasing [Robertson's] remarks, [Tate had] left the word "childish" which, of course, called for a rather emotional response from Silberstein.'[42] Indeed, Silberstein was incensed by the tone of the report and strongly criticized the anonymous referee for having been 'shamelessly arrogant and perfectly unfair'.[43]

In the early years of the 1930s the request of the editor to the referee came through a rather informal letter, sometimes carrying the editor's strong personal opinions about the value of the manuscript. An example of this attitude can be found in a letter in which Tate asked for suggestions about a paper on projective relativity by the theoretical physicist Banesh Hoffman. In the letter, *Physical Review*'s editor stated: 'I can make nothing of this paper and I get the impression that it is mostly throwing one's weight around.'[44] The requests concerning papers or letters to the editor that were not judged worthy of immediate publication were often very terse with direct questions such as, 'What shall I do with the enclosed letter to the Editor…?'[45] More often than not, in his request, the editor gave the impression of handing over the entire responsibility for the final decision about a submitted manuscript to the referee.

The referee's responses were similarly an expression of this sort of familiarity between the editors and the referee. The first referee reports prepared by Robertson were in the form of a simple letter, containing judgements about the paper and often including jokes.[46] Only on rare occasions did Robertson send in a separate set of sheets with a more formal report, including mathematical demonstrations.[47] Another example was provided by John C. Slater, who, in a letter to Tate assessing various issues, destroyed *en passant* a manuscript submitted to *Physical Review* by stating: 'The paper which you sent me to review seems to be entirely crazy. I presume you came to the same conclusions.'[48]

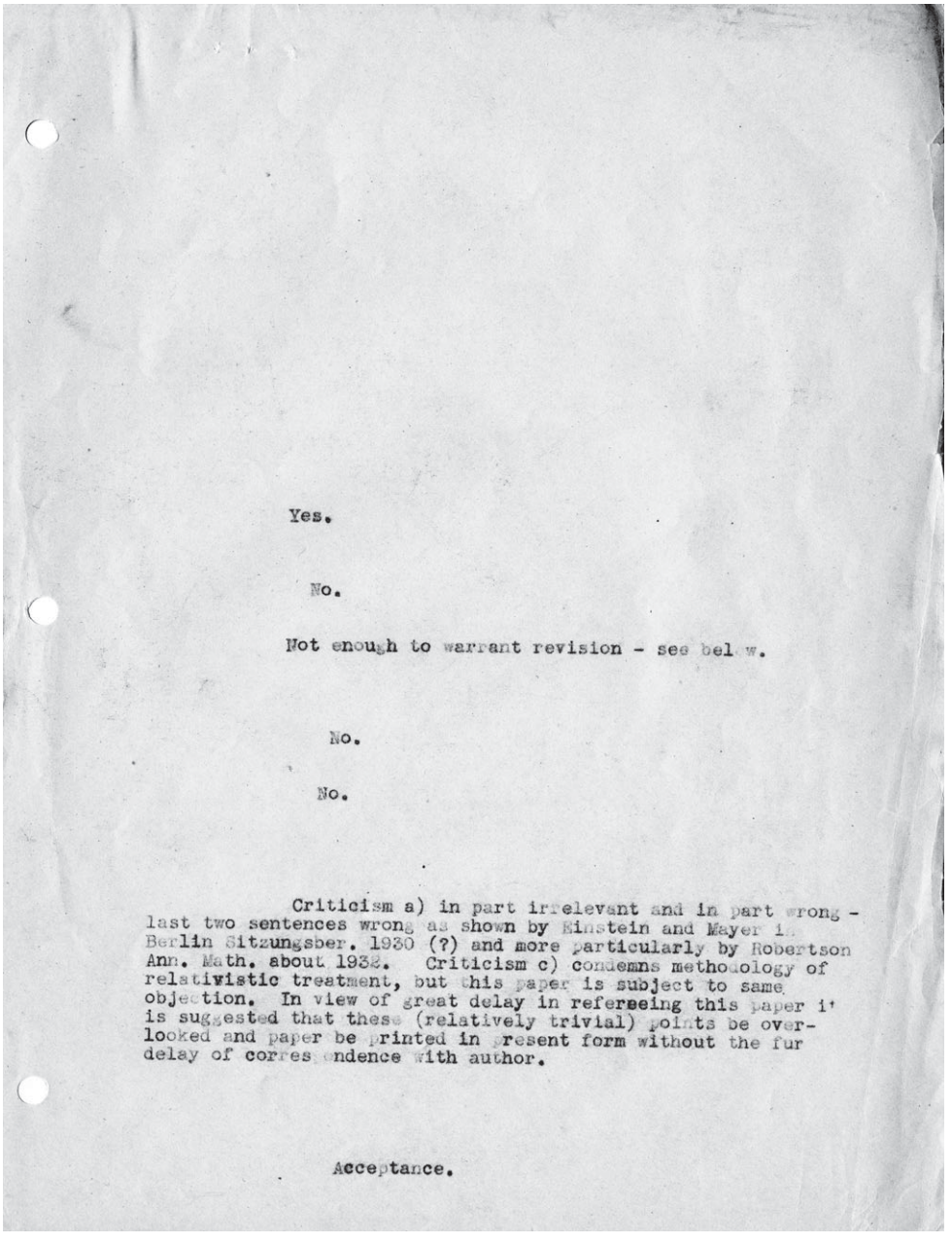
Although the trajectory is far from being linear, the correspondence shows that, starting from the mid-1930s, there were clear attempts to introduce a higher degree of formality in the

refereeing process. The assistant editor began sending as many letters as the editor-in-chief, if not more. In addition, the tone of editors' letters tended to be much more neutral than before—at times with only one question directly asking for an evaluation of the manuscript.[49] The increasing number of submitted papers probably led Tate to give up his attempt to read all the materials and in all probability he began forwarding some papers directly to the members of the editorial board and the experts of his circle of referees who, in his opinion, could judge the content of the papers.

More tellingly, in about 1935 a questionnaire was introduced to guide the referees' evaluation. From that moment onwards, an increasing number of Robertson's replies were based on this questionnaire. Figure 1 represents the first of these questionnaire-based reports about a manuscript submitted by the Swiss astronomer Fritz Zwicky—an image that only partly discloses the structure of the questionnaire.[50] A more complete document found in a different archive sheds light on its exact structure, though. The bulk of the questionnaire was divided into two three-question parts. The first asked for an evaluation of the content, and the second required from the referee an attentive judgement concerning the form of the manuscript. The very first question asked for a judgement about the scientific value of the paper by questioning whether the results contained therein were novel and interesting enough to warrant publication in the journal. The second content-related question addressed the issues concerning the mathematical and logical reasoning, and the third invited the referee to reflect on the novelty and importance of specific parts of the manuscript, to evaluate whether it could possibly be shortened. The form-related part required the referee to evaluate the order of presentation, the overall length of the paper and the English language employed by the author. The questionnaire ended by asking for 'other remarks' and 'recommendation'.[51]

It is not possible to isolate a clear-cut historical passage from the informal letter to the questionnaire as the method by which the editors requested an evaluation of the submitted manuscripts. Some of Robertson's referee reports sent after 1935 continued to be unrelated to any kind of questionnaire, and the historical process concerning the decision to employ it remains unclear.[52] In any case, it is evident that the questionnaire became more and more established as a way of increasing the effectiveness of the refereeing process. The number of Robertson's reports that were based on the questionnaire tended to increase rapidly over time and it remained a common procedure at least until the early 1950s.[53]

Although the process leading to the decision of introducing the questionnaire remains unknown, the questions contained therein make the concerns behind this decision apparent. The growth of the editorial activity required a more systematic organization in the handling of submitted papers. The questionnaire-based reports contained precise information including the title of the submitted paper and the dates of reception and reviewing, which permitted easier cataloguing of the increasing quantity of material.[54] Moreover, the length of papers began to be a major concern, as is evident from the fact that suggestions for reducing the length of the submitted papers were asked in both the content-related and form-related sections of the questionnaire. The stress on the issue of length is most probably an outcome of the tension between the increasing number of submitted manuscripts and the limited number of pages that the journal could contain. The questions concerning the English language might be a result of an increasing internationalization of the journal or, more probably, of the transformation of the American physics community, which had recently increased by admitting many German-speaking Jewish refugees.



Yes.

No.

Not enough to warrant revision - see below.

No.

No.

Criticism a) in part irrelevant and in part wrong - last two sentences wrong as shown by Einstein and Mayer in Berlin Sitzungsber. 1930 (?) and more particularly by Robertson Ann. Math. about 1932. Criticism c) condemns methodology of relativistic treatment, but this paper is subject to same objection. In view of great delay in refereeing this paper it is suggested that these (relatively trivial) points be overlooked and paper be printed in present form without the further delay of correspondence with author.

Acceptance.

Figure 1. Howard P. Robertson, referee report on 'Remarks on the redshift from nebulae' by F. Zwicky, *ca.* September 1935. (Howard Percy Robertson Papers, 10024-MS, Caltech Archives, California Institute of Technology, folder 7.12; courtesy Caltech Archives.)

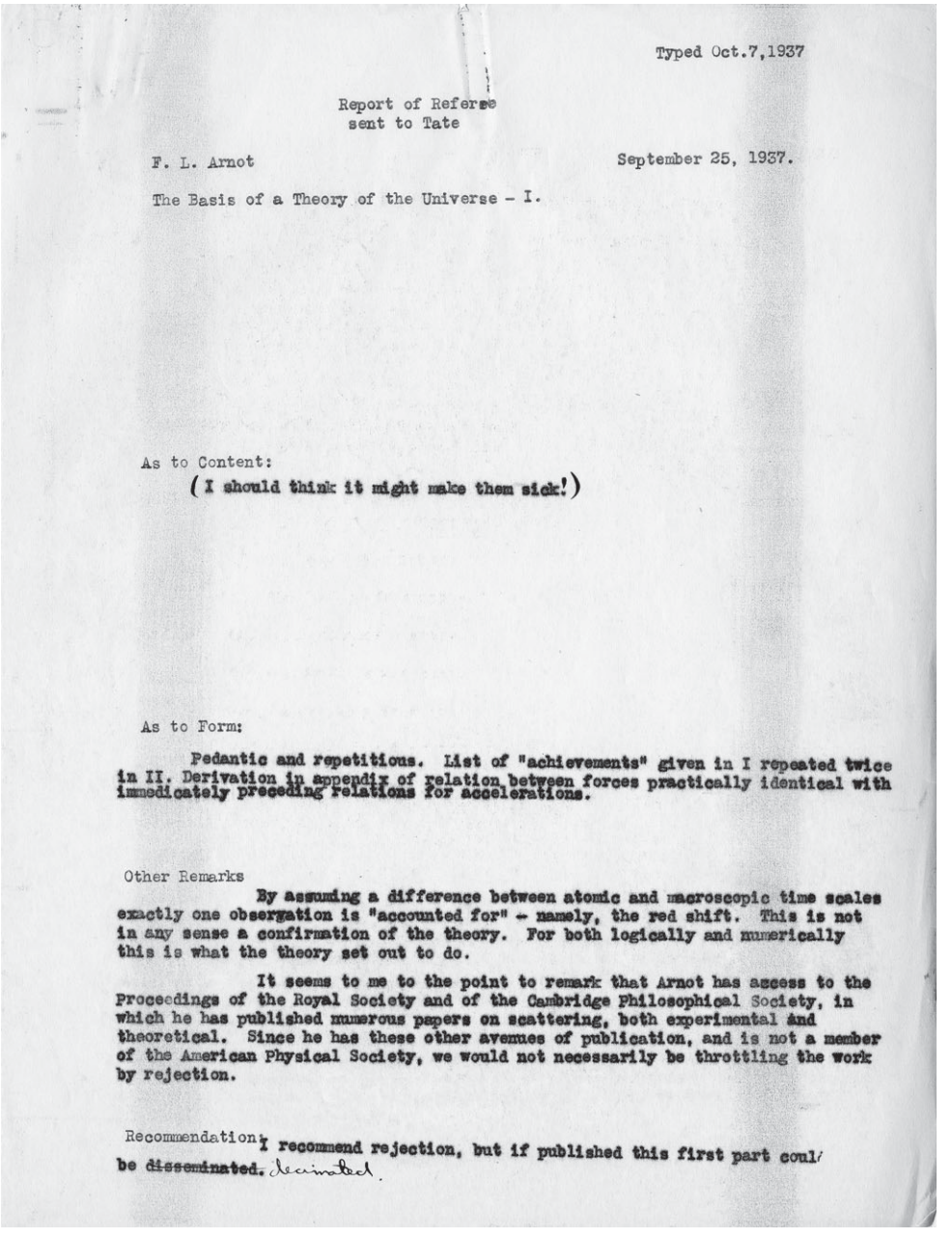
Whatever the reason for the introduction of the questionnaire, it had the potential to change the practice of refereeing in depth. The formulation of precise questions to guide the referee's evaluation of the manuscript could in principle impose some constraints on

the range of action of the referee. The questionnaire could, for instance, serve as a barrier to the more familiar relationships that the referees and the editors could have expressed more easily in personal letters. Moreover, the questionnaire asked the referee to focus on some topics (such as the order and length of the paper as well as the English language) that could have escaped the attention of a referee more interested in the scientific content of the submitted manuscript. The questionnaire was implicitly requesting the referee to modify his practice. No longer was his role limited to giving judgement about whether a paper was publishable or not. The referee was explicitly asked to improve the form of the presentation of the manuscript and then to suggest many more modifications than had been done before. The questionnaire could be a turn towards a standardization of the practice that entailed a transformation of the role of the referee from that of external judge to being similar to an advisor of the final publication—a role that has become quite common in today's practice.

The case of Robertson shows, however, that the informal-to-formal transformation implicitly embodied in the questionnaire did not occur easily. The way in which he handled the questionnaire shows that he did not make full use of all of the questions. Robertson's archived referee reports, mostly containing only Robertson's responses, could easily lead to the wrong impression that the questionnaire changed quite often, whereas a more detailed analysis shows that the changes depended on the unsystematic way in which Robertson employed it. Figure 2 shows an example of this partial employment of the questionnaire by Robertson. In the referee report he used the two parts of the questionnaire but completely ignored the actual questions (figure 2).[55] Moreover, Robertson often employed only the section 'remarks' for writing his judgements, thus reducing the potentially formalizing impact of the questionnaire on the refereeing practice itself.[56]

The frequent employment of the questionnaire did not lead to a change of Robertson's informal style in providing harsh judgements either. For example, as late as 1949 Robertson wrote a tongue-in-cheek questionnaire-based referee report for a paper written by the Portuguese mathematical physicist Antonio Gião, which Robertson did not consider worthy of publication. To the question asking whether there were ways in which to improve the order of presentation, Robertson answered: 'If it were written in invisible ink.'[57] The same holds true for a referee report about the Lorentzian interpretation of the gravitation equation put forward by the highly respected industrial physicist and disbeliever of relativity Herbert E. Ives. In his report Robertson ironically emphasized that 'the insistence on sticking to "Lorentzian" instead of the more appropriate "Minkowski" forces [was] a bit (ca. 40 yrs) out of date.'[58] Another example is to be found in the section 'remarks' of a referee report written in 1937. Robertson writes: 'this is the most wrong paper it has ever been my misfortune to law [*sic*] eyes upon', before concluding the report with the final recommendation: 'absolute and unconditional rejection, pointing out the numerous errors of his ways'.[59] It seems that Robertson could never resist including jokes or witty remarks in his reports. In November 1937 Buchta sent him a paper by the Dutch physicist Wander Johannes de Haas that had previously been rejected, asking whether they might be more generous. Robertson again replied in the negative, suggesting instead that 'the author ought to quit patting himself on the back.'[60]

Although Tate and Buchta tried to establish a formal procedure, possibly aimed at fostering an effective system for the evaluation of the proposed manuscripts, the familiarity between the editor and the chosen referees changed neither suddenly nor dramatically after the establishment of the new method. Until the end of the 1950s,



Typed Oct.7,1937

Report of Referee
sent to Tate

F. L. Arnot September 25, 1937.

The Basis of a Theory of the Universe - I.

As to Content:

(I should think it might make them sick!)

As to Form:

Pedantic and repetitious. List of "achievements" given in I repeated twice in II. Derivation in appendix of relation between forces practically identical with immediately preceding relations for accelerations.

Other Remarks

By assuming a difference between atomic and macroscopic time scales exactly one observation is "accounted for" - namely, the red shift. This is not in any sense a confirmation of the theory. For both logically and numerically this is what the theory set out to do.

It seems to me to the point to remark that Arnot has access to the Proceedings of the Royal Society and of the Cambridge Philosophical Society, in which he has published numerous papers on scattering, both experimental and theoretical. Since he has these other avenues of publication, and is not a member of the American Physical Society, we would not necessarily be throttling the work by rejection.

Recommendation: I recommend rejection, but if published this first part could be ~~disseminated~~. decimated.

Figure 2. Howard P. Robertson, referee report on 'The basis of a theory of the Universe. I' by F. L. Arnot, 25 September 1937 (Howard Percy Robertson Papers, 10024-MS, Caltech Archives, California Institute of Technology, folder 7.13; courtesy Caltech Archives.)

Robertson's style remained essentially sarcastic, especially, as one might well expect, with regard to those papers he did not consider worthy of publication in *Physical Review* or in any other scientific journal. The informal and ironic manner that Robertson employed in his

referee reports, moreover, did not depend on a possibly close relationship between Tate and himself but remained virtually unchanged if the request had originally come from Buchta or from subsequent editors.[61] The ways in which Robertson framed his views were an integral part of his refereeing. Although Robertson's ironical use of the questionnaire was plausibly exceptional, the difficulty in sticking to the new requirements embodied by the questionnaire was not peculiar to him. For instance, the Anglo-American theoretical physicist William F. G. Swann also found it difficult to use the questionnaire as a basis for the referee report and preferred to provide a separate letter in which he focused only on the content of the reviewed paper without addressing the formal aspects at all.[62]

The shaping of a referee: tensions between epistemic virtues and intellectual agendas

Besides shedding light on the actual practices, the correspondence between *Physical Review*'s editors and Robertson allows a series of elements concerning the perception of the refereeing activities according to the participants' categories at that time to be brought out. These elements include the criteria that were used to evaluate manuscripts, as well as the personal concerns expressed by the scientists grappling with the loosely defined activity of refereeing. In the present analysis I connect these elements to the attempts to define the epistemic virtues that a good referee should possess—namely, those intellectual dispositions that could serve as guiding principles for the referees' actions in the pursuit of the advancement of knowledge—and to the tensions related to these attempts.[63]

Many of the concerns exposed in the investigated correspondence were in part motivated by the fact that in the 1930s the practice was not completely accepted by all members of the physics community. It is well known that after Albert Einstein received a careful referee report contesting some of the conclusions of a paper he wrote with his collaborator Nathan Rosen, Einstein replied with an incensed letter stating that he had not authorized Tate to show his paper to other specialists before publication. Einstein's was only the most well-known example of the tension created by the different conceptions of the editorial process that was then current among practitioners. Many theoretical physicists coming from Germany or German-speaking countries, for instance, were unfamiliar with the practice of refereeing. The editors of authoritative German physics journals, including *Annalen der Physik*, did not usually seek external advice, and the publication of new papers by established scholars was almost automatic.[64]

In the unstable context in which refereeing was still perceived with suspicion by a part of the changing American physics community, the referees felt a strong responsibility to define the criteria that should be employed, as well as the epistemic virtues that a referee should possess to make the practice acceptable to the entire community. In this sense the referee was not only evaluating a paper but also implicitly building a complex entity—a *scientific persona*, as elaborated by Daston and Sibum—who had to possess specific epistemic virtues to embody the evolving function and its role in the production of certified knowledge.[65] At the same time, the investigation of these concerns reveals how the participants themselves perceived the function of the journal in that precise historic moment.

Arguably, the editor used several different criteria both at the beginning of the process, when he had to decide which papers were to be published and which were to be put under the judgement of a referee, and at the end, when he had to close a triangular

negotiation with a final decision about the editorial disposition of the manuscript. For their part the referees also employed different criteria in providing their assessments and recommendations. The criteria that Robertson employed might be separated into two kinds: one was general, applicable (and presumably applied) to every submitted paper; the other was specific to the subject matter of Robertson's expertise.

As for the general elements shaping the judgements and concerns of both the referee and the editor, if a paper was to be accepted, the mathematics employed had to be correct, and the reasoning exposed had to be consistent.[66] If this was the case, the paper was considered publishable, if it contained what the editor defined as 'new results' in the questionnaire. According to the participants' categories, this was the central element in deciding the fate of a manuscript.[67] The focus on the novelty of a paper's results had several implications. One of the most significant was that it specified the main attribute that a good referee should possess: he had to have an extended knowledge of the literature about a particular field to recognize whether similar findings had already been published. This quality played a major part when the editors were called to choose a new member of the editorial board, as shown in a letter by Tate on the appointment of Hans Bethe: 'I [nominated Bethe] for purely practical purposes. ... He has easily more encyclopedic information in his head on both the theory and the experiment in nuclear physics than almost any one else.'[68] This in turn implied that a good referee had to be the type of scientist who spent time carefully studying both past and current literature and who tried to be aware of all the work done in his field of expertise. This requirement was not trivial. Not every outstanding scientist of the period was eager to spend much time on the literature. Some important practitioners, including Einstein, had very different perceptions about the epistemic virtues of scientists: Einstein preferred to do all his derivations alone instead of relying on previous work by others.[69]

The focus on the novelty of results as a main motivation for publishing a certain research endeavour is also very instructive in understanding how the participants perceived the role of the journal in their discipline's dynamics and in the life of their specific community. Indeed, it suggests that the allocation of credit was commonly perceived to be one of the most important purposes, if not the most important purpose, of scientific journals.[70] Although this is not surprising, it is striking to note how often and to what degree this topic was central to the argument of the referee. Robertson frequently suggested that the paper should be transformed into a letter to the editor. In these cases, he disagreed with many of the passages actually contained within the paper but wished to avoid the possibility that an original result did not receive due credit because of his negative judgement.[71] In contrast, if the main result of a submitted paper had already been published as an abstract, he usually suggested rejection on the grounds that the abstract would have been sufficient to obtain credit for the finding.[72]

However, the novelty argument was occasionally used as a rhetorical device to recommend the rejection of particular approaches or theories that Robertson strongly disapproved of, even when they had only been published in non-English-language journals.[73] This case, in turn, suggests that Robertson did not consider the journal as a means of spreading knowledge quickly. He seemed to consider publication in *Physical Review* as a reward to the author rather than as a method for the readers to increase their knowledge. In this sense, *Physical Review* was seen as a way of allocating authority and power within a specific community rather than as a way of making knowledge products more readily available to a larger set of practitioners.

As confirmation of this view of the journal's function, the quality of previous achievements of the submitting author also played a major part in pondering whether a paper was to be published or rejected. Robertson was enough of an expert to spot possible mistakes, undue assumptions or non-physical consequences of a mathematical theory. However, more than once he relied on his personal judgement about the overall skill of the author to come to a final decision. After an exchange between Tate, Robertson and the author of a manuscript that Robertson did not consider worthy of publication because of its unclear physical significance, Robertson clearly stated that a scientist 'of good scientific standing ... has more right to be heard than any single referee has to throttle!'[74] On this note, when Robertson gave his assessment of a paper written by Synge, his judgement was deeply affected by Synge's prestige in the field. Robertson disagreed with the physical assumptions that Synge used in the manuscript and believed that in view of the observational data the proposed theory would be untenable in higher approximations. Nevertheless, Robertson asked Tate to dismiss his judgement and to ask someone else because Synge was a competent mathematician and his work was most probably 'correct to the approximations employed'.[75]

More prosaically, one of the most used criteria dealt with the affiliation of the author to the society of which the journal was the main publication venue.[76] When providing his judgements about the consistency of the physical assumptions employed, Robertson clearly differentiated between papers coming from authors who were members of the APS and those written by physicists who were not affiliated with the organization. This was true of a two-part paper submitted by the physicist Frederick L. Arnot entitled 'The basis of a theory of the Universe'. Because he found Arnot's approach strongly objectionable, Robertson recommended that both parts of the paper be rejected on the grounds that Arnot was not a member of the APS and that he seemed to have easy access to English journals in which he had previously published.[77] If an author was both an authority in his field and an esteemed member of the APS with a recognized status within the society (such as the fellowship), the referee did not feel the need to propose a rejection of his manuscripts, in spite of any negative judgements about the scientific cogency of the work submitted. Robertson exposed this feeling in his recommendation of a two-part paper by the well-known German–American physicist Alfred Landé in 1939. He wrote: 'Landé is a Fellow of the [APS], and enjoys a good reputation in physics circles. As such, he has the right to be heard.'[78]

Besides the above-mentioned criteria, which could be applied to any submitted paper, a fundamental role was played by epistemological considerations largely related to the unstable position of Robertson's field of expertise in the scientific landscape. The theory of general relativity and related research posed several problems arising from the undefined and shifting disciplinary boundaries between mathematics and physics, and the critical role of theoretical assumptions in such a borderline field. Because in his own scientific activities, especially in his approach to relativistic cosmology, Robertson held very clear opinions about the relations between empirical observations, physical assumptions and mathematical reasoning, in his refereeing Robertson did not disdain taking a definitive stance against the publication of papers that presented approaches that he either did not consider promising or regarded as being too close to a rationalistic approach to theoretical physics and cosmology.[79] In the 1930s the main targets of Robertson's public criticisms were the special relativistic cosmological theory put forward by Edward A. Milne and Eddington's fundamental theory unifying general relativity, quantum theory and cosmology, based on numerological reasoning about the dimensionless ratios of universal constants. In his view,

all these attempts were driven by rationalistic perspectives, which according to him were very far removed from the actual way in which physical theories could, and actually did, progress. In his correspondence Robertson continuously criticized the above-mentioned approaches and he came to identify a sort of English school of cosmological thought against which he was constantly striving.[80]

Robertson's personal scientific battle repeatedly affected his work as a referee. His opposition to a particular approach played a prominent part in his final judgement as to whether a paper was to be published or rejected. If the physical hypotheses did not satisfy his personal criteria about the soundness of the assumptions, he regarded the hypotheses as *ad hoc* and tended to suggest the rejection of papers that employed reasoning he considered similar to those employed by the above-mentioned English savants. In the same vein he promoted his own view by suggesting the publication of papers explicitly opposing these views. The final recommendation of a paper submitted by Boris Podolsky that aimed at criticizing a previous work by Eddington does not leave much doubt about Robertson's explicit motivation in evaluating the paper. Robertson wrote: 'Not too exciting, but it seems to me justifiable to publish refutations of stuff published by supposedly reputable scientists as A.S.E. [Arthur S. Eddington].'[81] Moreover, he often expressed judgements as to whether a paper was a promising attack on current problems. Even when he found no error, he considered himself in the position to evaluate which research projects might be fruitful enough for further research, thus in effect employing his own general approach to judge the future developments of the fields in which he was actively involved.[82]

The arguments based on epistemological considerations strongly contrasted with the one that saw novelty of results as the main criterion in evaluating whether a paper was worthy of publication in *Physical Review*. In employing these different criteria, Robertson was actually referring to different views of the journal's function. When epistemological considerations played the major part, Robertson seemed to consider the journal as a place to 'immortalize', or make eternal, the research results of a colleague.[83] Although Robertson's own expression 'immortalize' can be interpreted in different ways, it seems that Robertson intended, again, the journal to be a place where the authors' scientific results would be rewarded rather than made available to a broader readership. But in this case, the question that plausibly was implicit in Robertson's evaluation was: what kind of writing was worth receiving the status of being read by virtually the entire physics community?

Although both criteria were eventually understood in the light of what the journal provided for the author more than for the reader, the criterion of epistemological soundness related to the function of scientific journals as 'immortalization' was in conflict with the criteria based on the view of the journal as a place to allocate recognition for novel discoveries. This contrast emerged explicitly in some of Robertson's referee reports, in which he exposed his concerns about the criteria for the acceptance or rejection of submitted manuscripts. Reflecting on a paper written by Haas, Robertson expressed his opposition to the idea that a future incorporation of what he considered Haas's *ad hoc* hypothesis in a more general field theory should lead to Haas's being credited for having proposed the idea as an 'isolated hunch'.[84] In using different criteria for assessing the validity of a manuscript depending on varying views of the journal's role and on loosely defined images of the epistemic virtues that a good referee should possess, Robertson made some original syntheses, which in general tended towards the promotion of his own scientific agenda. In turn, this resulted in the view that the referee had both the right and the power to indicate what direction future research should take.

Conclusion

In the 1930s, *Physical Review* was establishing a form of refereeing system under the assumption that this practice would eventually be accepted by all members of the American physics community. In implementing the practice, both Tate and Robertson followed an implicit agenda. The former wished to improve the qualitative standards of the journal; the latter wanted to improve the status of general relativity and, more specifically, of relativistic cosmology within the physics community. As I have shown above, when a paper was sent to the referee, the editor tended to follow the referee's judgement, although with some degree of arbitration. But the referee used various and contrasting criteria to judge the validity of the manuscripts. Because of the particular epistemic status of general relativity theory in that period, one of the criteria dealt with epistemological considerations about what Robertson believed to be unduly physical assumptions. The balance between the different criteria about the validity of the submitted manuscript was very complex and created a certain tension, which Robertson's referee reports sometimes made explicit.

The contrast between the criteria employed can be understood on three different levels. On the most basic level, the tension concerned the difficulties in establishing a clear scale of values that could be followed in every case. Robertson tended to change this scale of criteria depending on the general programme of which the paper was an expression. Although novelty of results seemed to be the fundamental point in most referee reports, it might also be used rhetorically to avoid the publication of papers that Robertson found objectionable according to his own epistemological views. More subtly, novelty was turned into a negative feature when, as in Haas's manuscript, it could become a way to allocate credit when Robertson did not consider the approach worthy of that credit. In this sense, even the criterion of novelty was rather flexible, and at times of secondary importance with regard to Robertson's more general view about the soundness and relevance of the proposed manuscript.

At the level of the journal's function according to the participants' categories, the tension emphasizes that those involved were continuously reconfiguring the meaning of a scientific journal. Robertson did not seem to consider fundamental problems such as the prompt and rapid diffusion of scientific knowledge. He was always concerned more with what the journal could do for the author than with what it could do for the reader. In this sense, the tension between the different views of the journal's function was not so strong and these views could be negotiated and changed according to the preference of the referee.

At the third level, the epistemic virtues of the referees were also negotiated. A fundamental requirement was that the referee had an ample awareness of the literature and would be willing to take time from his own research to serve the community. This suggests a precise kind of *scientific persona* who was not particularly concerned with geniality or competitiveness but should rather be a repository of common knowledge. Moreover, Robertson believed that he had to be open to the publication of pluralistic views. But even in this case this perception of the role of the referee contrasted with his own scientific needs to promote his own approach. In all the cases in which Robertson had to evaluate approaches that he constantly challenged in his daily scientific activity, his own agenda became more relevant than the virtues he ideally considered important in acting as a good referee.

If we generalize these points, we can conclude that in the 1930s the referee had enormous power to prevent the publication of papers from scholars whose overall approach he had direct quarrels about. Because the editor tended to follow the advice of the only referee called to evaluate the manuscript, it seems that the refereeing practice of the time could

hardly be called peer review if the egalitarian meaning of the word 'peer' in the Mertonian sense is to be maintained. The referee was undoubtedly an authority in the field and had an enormous degree of power—a power that was realized in several ways, from the dismissive language employed in the referee reports to the familiarity with the editor. More importantly, the referee felt that he had the right to define what was a meaningful or promising direction for future research, along with the tendency to support his own agenda and epistemological convictions.

The link between the present research and previous studies concerning the postwar period suggests that the evolution of the practice of refereeing in *Physical Review* was disrupted, particularly by the war and its related social changes and by problems concerning the disclosure of scientific results that may have been employed in military research. Tate's death in 1950 also posed a sort of caesura; as shown by Kaiser, Samuel Goudsmit—the editor of *Physical Review* from 1951—and his assistant editor, Simon Pasternak, reinvented ways of implementing the refereeing practice, thus implying that the editorial evolution was also related to the contingent persons with the responsibility to implement the practice. They employed memos to their colleagues at the Brookhaven National Laboratory, who were asked to fill out the form to provide initial assessments and judgements about the submitted manuscripts. The procedure helped understand whether the decision about the paper could be made through in-house connections or whether an external referee was in fact needed. The various formalizing attempts evolved rapidly and in 1960 the editors began to send all papers to external referees, with a pre-printed referee form asking for comments that would be transmitted directly to the authors. The number of both submitted papers and referees increased further in the following years, requiring an automation of the entire process. In this historical process one can identify with Kaiser a development from a practice based on familiarity (close relationships between the editors and the referees) to an impersonal mechanical system, which was necessary to cope with the rapidly increasing number of submitted manuscripts. Kaiser argues that the institutionalization of a formal peer-review process for the evaluation of submitted manuscripts was mostly a consequence of, or at least was strongly related to, the enormous growth of the population of physicists and the related pace of specialization and information overload in the USA in the period after World War II and during the Cold War.[85] By taking into consideration a similar period (1948–56), Zuckerman and Merton have argued that during these years the evaluation pattern of *Physical Review* was close to what they called the 'model of expertise'.[86]

If we accept Zuckerman and Merton's categorization summarized in the introduction, one might say that in the 1930s the refereeing model followed by *Physical Review* was instead similar to what they called the 'oligarchical model' of the refereeing practice. The present conclusion seems to be in line with a recent study by Imogen Clarke, who argues that in the 1920s the British physics periodicals *Proceedings of the Royal Society of London* A and *Philosophical Magazine* were both strongly influenced by small networks of experts granted the authority to evaluate the research endeavours of all the other practitioners.[87] Given this transformation in the approaches to the refereeing process within *Physical Review* from the 1930s to the 1950s, the quite rapid historical development of the practice and its gradual transformation into what we now call peer review suggests, in agreement with Kaiser and Burnham, that this process might well have been related to an evolution mostly due to the quantitative growths of the number of specialists and to the associated pace of specialization.[88]

Acknowledgements

This research was supported by the Maurice A. Biot Grant-in-Aid of the California Institute of Technology (Caltech) Archives. I am deeply grateful to the archivists of the Caltech Archives for their support. I am indebted to Amy Halsted and to the editors of *Physical Review*, who permitted access to the editorial tables recording the paths of the manuscripts submitted from 1932 to 1937. I am also grateful to the librarians and archivists of the Max Planck Institute for the History of Science library, of the American Philosophical Society and of the Niels Bohr Library and Archives (American Institute of Physics). My thanks go to David Kaiser for having given me the opportunity to read a chapter of his forthcoming book. A previous version of this paper was presented at the conference 'Publish or perish? The past, present and future of the scientific periodical' held at the Royal Society on 19–21 March 2015. I heartily thank the organizers, Aileen Fyfe, Julie McDougall-Waters and Noah Moxham, and also the attendees, for useful discussions. I am especially indebted to Melinda Baldwin, two anonymous referees and the editor of this journal for their insightful comments.

Notes

1 For some discussion about the flaws inherent in peer review see, for example, A. C. Weller, *Editorial peer review: its strengths and weaknesses* (Information Today Inc., Medford, NJ, 2001); R. Smith, 'Peer review: a flawed process at the heart of science and journals', *J. R. Soc. Med.* **99**, 178–182 (2006), and D. Rennie, 'Editorial peer review: its development and rationale', in *Peer review in health sciences* (ed. F. Goodlee and T. Jefferson), pp. 1–13 (BMJ Books, London, 1999). In 2006 the website of *Nature* hosted a debate on peer review that presented very contrasting views: see http://www.nature.com/nature/peerreview/debate/ (accessed 23 January 2015).

2 See, for example, E. Cattaneo and G. Corbellin, 'Stem cells: taking a stand against pseudoscience', *Nature* **510**, 333–335 (2014); 'Paper on controversial stem cell "stamina therapy" retracted', *Retraction Watch*, available online at http://retractionwatch.com/2014/12/10/paper-controversial-stem-cell-stamina-therapy-retracted/ (accessed 11 March 2015); 'Deregulating stem cell quackery in Italy? A bad omen', *Scienceblogs*, posted 6 May 2013, available at http://scienceblogs.com/insolence/2013/05/06/deregulating-stem-cell-quackery-in-italy/ (accessed 11 March 2015); J. Travis and L. Margottini, 'More on quake, radongate, Guiliani', *Science News*, posted 7 April 2009, available at http://news.sciencemag.org/2009/04/more-quake-radongate-guiliani (accessed 11 March 2015).

3 A series of essays criticizing the conceptual unity of science and the methodological value of unification in science is included in P. Galison and D. J. Stump (eds) *The disunity of science: boundaries, contexts, and power* (Stanford University Press, 1996). For a discussion of the methodological pluralism of the different sciences see especially I. Hacking, 'The disunities of the sciences', *ibid.*, pp. 37–74.

4 The relevance of peer review in certifying that a particular result is scientific is given at times extreme emphasis in science journalism. See, for example, S. Anthony, 'CERN's Higgs boson discovery passes peer review, becomes actual science', *ExtremeTech* (10 September 2012), available at http://www.extremetech.com/extreme/135756-cerns-higgs-boson-discovery-passes-peer-review-becomes-actual-science (accessed 11 March 2015).

5 For a sociological study on the boundary work and the employment of peer review in defining the boundaries of science see T. F. Gieryn, *Cultural boundaries of science: credibility on the line* (University of Chicago Press, 1999), esp. ch. 4; for an interesting case study about the tension

between academic freedom and peer-review-based boundary work see N. Nattrass, 'Defending the boundaries of science: AIDS denialism, peer review and the medical hypotheses saga', *Sociol. Health Illn.* **33**, 507–521 (2011).

6 The expression 'standard model of scientific periodical' was introduced by A. Johns at the plenary roundtable discussion opening the conference 'Publish or perish? The past, present and future of the scientific periodical' held at the Royal Society on 19–21 March 2015. A selection of the talks delivered at the conference has been published in A. Fyfe, J. McDougall-Waters, N. Moxham (eds), '350 years of scientific periodicals', *Notes Rec.* **69**, 227–352 (2015).

7 For examples and a thorough analysis of contemporary fringe science see M. Gordin, *The pseudoscience wars: Immanuel Velikovsky and the birth of the modern fringe* (University of Chicago Press, 2012), esp. pp. 195–212; a further example is the journal *APEIRON: Studies in Infinite Nature*, which was dedicated to publishing unorthodox physical theories. See http://redshift.vif.com/Apeiron_Home.htm (accessed 6 January 2016).

8 See Weller, *op. cit.* (note 1); T. H. P. Gould, *Do we still need peer review? An argument for change* (Scarecrow Press, Lanham, MD, 2013); J. J. Gaudet, 'All that glitters is not gold: the shaping of contemporary journal peer review at scientific and medical journals', *uO Res.* (2014), p. 4 (available at http://hdl.handle.net/10393/31238; accessed 6 January 2016).

9 Gaudet, *op. cit.* (note 8), p. 2.

10 In the entry 'peer review' of the *Oxford English Dictionary* the first example of the expression 'peer review system' dates back to 1967, in an article from *Washington Post* discussing a process of evaluation by review boards in the medical system. The first use in the academic sense dates back to 1971 and it is again related to medical practices. See http://www.oed.com/view/Entry/139736?rskey=1EeMPZ&result=1#eid (accessed 1 September 2015; requires subscription).

11 H. Zuckerman and R. Merton, 'Patterns of evaluation in science: institutionalization, structure and functions of the referee system', *Minerva* **9**, 66–100 (1971), at p. 88.

12 For a historical analysis of the establishment of the refereeing practices in royal academies as an evolution of book censorship see M. Biagioli, 'From book censorship to academic peer review', *Emergences* **12**, 11–45 (2001). A description of the historical evolution of the refereeing practices is in J. Burnham, 'The evolution of editorial peer review', *J. Am. Med. Assoc.* **263**, 1323–1329 (1990), and D. Kronick, 'Peer review in 18th-century scientific journalism', *J. Am. Med. Assoc.* **263**, 1321–1322 (1990). More recently, several important steps have been made in the understanding of this complex historical process: Alex Csiszar, 'Broken pieces of fact: the scientific periodical and the politics of search in nineteenth-century France and Britain', PhD thesis, Harvard University (2010); I. Clarke, 'The gatekeepers of modern physics: periodicals and peer review in 1920s Britain', *Isis* **106**, 70–93 (2015); I. Clarke and J. Mussell, 'Conservative attitudes to old-established organs: Oliver Lodge and *Philosophical Magazine*', *Notes Rec.* **69**, 321–336 (2015); M. Baldwin, 'Credibility, peer review, and *Nature*, 1945–1990', *Notes Rec.* **69**, 337–352 (2015); A. Fyfe, J. McDougall-Waters and N. Moxham, 'Guest editorial: 350 years of scientific periodicals', *Notes Rec.* **69**, 227–239 (2015); J. McDougall-Waters, 'Refereeing in the *Philosophical Transactions*, 1830–1900', presented at the conference 'Publish or perish? The past, present and future of the scientific periodical'. The topic is also at the centre of ongoing research projects such as Melinda Baldwin's project on the establishment of peer review in the twentieth century and the project 'Publishing the *Philosophical Transactions*' funded by the Arts and Humanities Research Council.

13 For the concept of big picture see J. Secord, 'Introduction', *Br. J. Hist. Sci.* **26**, 387–389 (1993).

14 The dominant role of the journal *Physical Review* in the physics landscape is argued in Zuckerman and Merton, *op. cit.* (note 11). The deployment of the tool Science Citation Index since 1963 has made even more evident the dominant position of the journal. See Gaudet, *op. cit.* (note 8), p. 4.

15 D. Kaiser, 'Information overload', in D. Kaiser, *American physics and the Cold War bubble* (University of Chicago Press, in the press), ch. 3.
16 For the concept of scientific persona and its historical implications see L. Daston and H. O. Sibum, 'Introduction: scientific personae and their histories', *Sci. Context* **16**, 1–8 (2003); for the definition of epistemic virtues I refer to L. Daston and P. Galison, *Objectivity* (Zone Books, New York, 2007), esp. pp. 39–42.
17 J. Eisenstaedt, 'La relativité générale à l'étiage: 1925–1955', *Arch. Hist. Exact Sci.* **35**, 115–185 (1986); *idem*, 'The low water mark of general relativity: 1925–1955', in *Einstein and the history of general relativity* (*Einstein studies*, vol. 1) (ed. D. Howard and J. Stachel), pp. 277–292 (Birkhäuser, Boston, MA, 1987).
18 See A. Blum, R. Lalli and J. Renn, 'The re-invention of general relativity: a historiographical framework for assessing 100 years of curved space-time', *Isis* **106**, 598–620 (2015).
19 A. Taub, 'Obituary: Prof. H. P. Robertson', *Nature* **192**, 797–798 (1961); J. Greenstein, 'Howard Percy Robertson, 1903–1961', *Biogr. Mems Natl Acad. Sci.* **51**, 343–364 (1980). On the status of theoretical physics in the USA before World War II see S. Schweber, 'The empiricist Temper Regnant: theoretical physics in the United States 1920–1950', *Hist. Stud. Phys. Biol. Sci.* **17**, 55–98 (1986).
20 J. T. Tate to H. P. Robertson, 5 February 1937, Howard Percy Robertson Papers (hereafter HRP), 10024-MS, Caltech Archives, California Institute of Technology, folder 7.13; J. W. Buchta to H. P. Robertson, 4 January 1940, HRP, folder 7.14.
21 J. L. Synge to H. P. Robertson, 27 April 1933, HRP, folder 5.25.
22 B. Bertotti, 'Cosmology, a peculiar science', in *Modern cosmology in retrospect* (ed. B. Bertotti, R. Balbinot, S. Bergia and A. Messina), pp. 3–9 (Cambridge University Press, 1990); H. Kragh, *Quantum generations: a history of physics in the twentieth century* (Princeton University Press, 1999), esp. pp. 349–365; G. Gale, 'Cosmology: methodological debates in the 1930s and 1940s', in *Stanford Encyclopedia of Philosophy* (spring 2014 edition) (ed. Edward N. Zalta), available at http://plato.stanford.edu/archives/spr2014/entries/cosmology-30s/; Blum *et al.*, *op. cit.* (note 18).
23 'Proceedings of the American Physical Society: minutes of the sixty-fourth meeting', *Phys. Rev.* **1**, 60–84 (1913), at p. 63.
24 P. Hartman, *A memoir on* The Physical Review*: a history of the first hundred years* (AIP Press, New York, 1994), pp. 9–10; D. J. Kevles, *The physicists: the history of a scientific community in modern America* (Harvard University Press, Cambridge, MA, 1995), at p. 282.
25 Kevles, *op. cit.* (note 24), esp. pp. 200–286; K. Sopka, *Quantum physics in America: the years through 1935* (American Institute of Physics, New York, 1988); Kragh, *op. cit.* (note 22).
26 Hartman, *op. cit.* (note 24), p. 106.
27 A. O. C. Nier and J. J. H. Van Vleck, 'John Torrence Tate, 1889–1950', *Biogr. Mems Natl Acad. Sci.* **47**, 461–484 (1975).
28 R. Lalli, 'A new scientific journal takes the scene: the birth of *Reviews of Modern Physics*', *Annln Phys.* **526**, A83–A87 (2014).
29 See M. Baldwin, '"Keeping in the race": physics, publication speed and national publishing strategies in *Nature*, 1895–1939', *Br. J. Hist. Sci.* **47**, 257–279 (2014).
30 See Hartman, *op. cit.* (note 24), and 'Finding Aid J. William Buchta Papers, 1923–1966: biographical sketch of J. William Buchta (1894–1966)', University of Minnesota Archives, available at http://special.lib.umn.edu/findaid/xml/uarc00367.xml (accessed 21 September 2015).
31 Hartman, *op. cit.* (note 24), p. 158.
32 *Physical Review* referee files, 1932–1943, Niels Bohr Library and Archive, American Institute of Physics.
33 Kaiser, *op. cit.* (note 15).
34 Robertson to Tate, 18 February 1937, HRP, folder 7.13.

35 Zuckerman and Merton, *op. cit.* (note 11), p. 78; a calculation of the acceptance rate in the period 1932–1937 shows that about 70% of the submitted papers were published in *Physical Review*.
36 See, for example, Robertson to Tate, 26 December 1931, HRP, folder 7.12.
37 Robertson to Tate, 9 June 1930, HRP, folder 7.12.
38 Robertson to Tate, 18 August 1930, HRP, folder 7.12.
39 Tate to Robertson, 30 January 1931, HRP, folder 7.12.
40 See, for example, Buchta to Robertson, 27 July 1939, HRP, folder 7.13.
41 H. P. Robertson, referee report on 'On the Quantization of Mass', by B. Podolsky and H. Branson, 11 July 1939, HRP, folder 7.13.
42 Tate to Robertson, 14 April 1936, HRP, folder 7.12.
43 Silberstein to Tate, 6 April 1936, HRP, folder 7.12.
44 Tate to Robertson, 7 January 1933, HRP, folder 7.12.
45 Tate to Robertson, 29 September 1936, HRP, folder 7.12.
46 See, for example, Robertson to Tate, 15 May 1933, HRP, folder 7.12.
47 An example of more impersonal referee reports is H. P. Robertson, referee report on Lanczos' 'A New Connection between Electricity and Riemannian Geometry', HRP, folder 7.12.
48 Slater to Tate, 2 October 1930, John Clarke Slater Papers, Mss.B.SL2p, American Philosophical Society, folder Tate/1.
49 Buchta to Robertson, 14 July 1938, HRP, folder 7.13.
50 H. P. Robertson, referee report on 'Remarks on the Redshift from Nebulae' by F. Zwicky, *ca.* September 1935, HRP, folder 7.12.
51 W. F. G. Swann, referee report on 'Dipole Origin of the Earth's Electric Gradient' by M. C. Holmes, 2 November 1936, William Francis Gray Swann Papers, Mss. B. Sw. 1, American Philosophical Society, folder John T. Tate 1934–1936.
52 See, for example, Robertson to Tate, 23 May 1936, HRP, folder 7.12.
53 The last of Robertson's referee reports based on the questionnaire is dated 18 September 1952. H. P. Robertson, referee report on 'Arbitrary Concentration of matter and the Schwarzschild Singularity', by Amalkumar Raychaudhury, HRP, folder 7.14.
54 See, for example, H. P. Robertson, referee report on 'Static Universe and Nebular Red Shift II', by S. Sambursky and M. Schiffer, 30 September 1937, HRP, folder 7.13; and H. P. Robertson, referee report on 'Stellar Statistics and Bode's Law', by H. Jehle, *ca*. June 1938, HRP, folder 7.13.
55 H. P. Robertson, referee report on 'The Basis of a Theory of the Universe I' by F. L. Arnot, 25 September 1937, HRP, folder 7.13.
56 H. P. Robertson, referee report on 'On the Interchange Energy of Two Free Electrons', by Podolsky and Ludeke, *ca.* 1938, HRP, folder 7.13.
57 H. P. Robertson, referee report on 'On the General Motion of Matter at a Cosmological Scale' by Antonio Gião, *ca.* December 1949, HRP, folder 7.14.
58 H. P. Robertson, referee report on 'The Physical Significance of Birkhoff's Gravitational Equations', by H. E. Ives, *ca.* March 1947, HRP, folder 7.14.
59 H. P. Robertson, referee report on 'On a Gravitational Theory of Neutrinos', by Gleb Wataghin, *ca.* February 1937, HRP, folder 7.13.
60 H. P. Robertson, referee report on 'A Relation Between the Gravitational Constant and Hubble's Factor', by W. J. Haas, November 1937, HRP, folder 7.13.
61 See Buchta to Robertson, 21 December 1949, HRP, folder 7.14; H. P. Robertson, *op. cit.* (note 53); and Buchta to Robertson, 3 January 1950, HRP, folder 7.14.
62 Swann to Tate, 5 November 1936, William Francis Gray Swann Papers, Mss. B. Sw. 1, American Philosophical Society, folder John T. Tate 1934–1936.
63 For the notion of epistemic virtues see Daston and Galison, *op. cit.* (note 16).
64 D. Kennefick, *Traveling at the speed of thought: Einstein and the quest for gravitational waves* (Princeton University Press, 2007), pp. 79–104; on the editorial practices followed in the

eminent *Annalen der Physik* in the first decades of the twentieth century see C. Jungnickel and R. McCormmach, *Intellectual mastery of nature*, vol. 2 (*The Now Mighty Theoretical Physics 1870–1925*) (University of Chicago Press, 1986), pp. 310–321; on Planck's editorial activity see D. Hoffmann, '"... you can't say to anyone to their face: your paper is rubbish.' Max Planck as editor of *Annalen Der Physik*', *Annln Phys.* **17**, 273–301 (2008).

65 Daston and Sibum, *op. cit.* (note 16).

66 See, for example, Robertson, *op. cit.* (note 41), and Tate to Robertson, 26 January 1949, HRP, folder 7.14.

67 On the relevance of novelty see, for example, H. P. Robertson, referee report on 'On Null-geodesics and Null-Corpuscles in the Theory of Relativity', by C. G. Pendse, 16 March 1939, HRP, folder 7.13. The relevance of concerns about the novelty of published results in the evolution of the formal peer-review system is indicated by the formulation of the famous Ingelfinger rule. In 1969 the editor of *New England Journal of Medicine*, Franz J. Ingelfinger, clearly stated that systematic reviewing of submitted manuscripts was essential to avoid the publication of unoriginal results. See F. Ingelfinger, 'Peer review in biomedical publication', *Am. J. Med.* **56**, 686–692 (1974).

68 Tate to Robertson, 7 January 1936, HRP, folder 7.12.

69 D. Kennefick, 'Einstein versus the *Physical Review*', *Phys. Today* **58**, 43–48 (2005), at p. 46.

70 Since the end of the nineteenth century the publication of scientific results had become the necessary way to claim credit for a novel discovery: see Csiszar, *op. cit.* (note 12), ch. 2. On the importance of this role in shaping the editorial policies of scientific periodicals see Baldwin, *op. cit.* (note 29).

71 H. P. Robertson, referee report on 'The Relativistic Theory of Spinning Charged Particles in Electromagnetic Fields', by William Band, *ca.* April 1939, HRP, folder 7.13.

72 H. P. Robertson, referee report on 'The Red Shift of the Nebular Spectrum Lines', by Ira Freeman, HRP, folder 7.13; see also Robertson, *op. cit.* (note 60), and Robertson to Goudsmit, 13 December 1954, HRP, folder 7.14.

73 Robertson, *op. cit.* (note 54).

74 Robertson to Tate, 27 July 1936, HRP, folder 7.12.

75 Robertson to Tate, 16 January 1935, HRP, folder 7.12.

76 This was often an important criterion for accepting papers in the publication venues of the various societies, because members assumed that they had the right to publish in their society's journals. See Burnham, *op. cit.* (note 12), esp. p. 1325. The belief that an author had this kind of right sparked a long controversy between the American astronomer and relativity disbeliever Charles L. Poor and the editor of *Science*, James McKeen Cattell. Because Poor was a member of the American Association for the Advancement of Science, he was convinced that it was his right to publish his criticisms about general relativity in the journal *Science*, the AAAS publication venue, but Cattell rejected Poor's papers after having received adverse opinions by the referees. See Poor to James McKeen Cattell, 6 May 1940, Johns Hopkins Library, Sheridan Library, Charles L. Poor Papers, folder Cattell.

77 Robertson, *op. cit.* (note 55), and H. P. Robertson, referee report on 'The Basis of a Theory of the Universe. II', by F. L. Arnot, HRP, folder 7.13.

78 H. P. Robertson, referee report on 'On the Ratio of $h/2$ and e^2/c' Parts I and II', by Alfred Landé, HRP, folder 7.13.

79 An analysis of the debates on cosmology and of their philosophical implications is in G. Gale and John Urani, 'Milne, Bondi and the "second way" to cosmology', in *The expanding worlds of general relativity* (*Einstein studies*, vol. 7) (ed. H. Goenner, J. Renn, J. Ritter and T. Sauer), pp. 343–375 (Birkhäuser, Boston, MA, 1999), and G. Gale, *op. cit.* (note 22).

80 See H. P. Robertson, referee report on 'The Basis of a Theory of the Universe. II', by F. L. Arnot, HRP, folder 7.13; see also Robertson to Eric Temple Bell, 15 September 1936, HRP, folder 1.13.

81 P. Robertson, referee report on ‘On Eddington’s Treatment of Dirac’s Equation’, by Boris Podolsky, 13 February 1938, HRP, folder 1.13.
82 See, for example, H. P. Robertson, ‘Referee Report on ‘Static Universe and Nebular Red Shift’ by S. Sambursky, 1 April 1937, HRP, folder 1.13.
83 Robertson to Tate, 22 May 1936, HRP, folder 1.12.
84 *Ibid*.
85 Kaiser, *op. cit.* (note 15).
86 Zuckerman and Merton, *op. cit.* (note 11), p. 88.
87 Clarke, *op. cit.* (note 12).
88 Kaiser, *op. cit.* (note 15), and Burnham, *op. cit.* (note 12).