Received February 2008
Revised December 2008
Accepted February 2009

An instrument for the self-appraisal of scientific research performance

James C. Ryan and Syed Awais Tipu
University of Sharjah, Sharjah, United Arab Emirates

Abstract

Purpose – The purpose of this paper is to present findings related to an instrument for the self-appraisal of scientists' research performance, and highlight the suitability of self-appraisal instruments for members of the scientific community.

Design/methodology/approach – An examination of the literature on self-appraisal and the measurement of scientific research is presented. The initial development of the instrument employed qualitative methods through interview and discussions with PhD-qualified scientific researchers ($n = 13$). A quantitative investigation of the usefulness of the instrument was then conducted on a sample of biological and chemical research scientists ($n = 270$). Results were compared with an existing performance measure and examined for representative reliability.

Findings – Results suggest that the instrument may be a reliable measure of research performance when used in a non-critical context.

Research limitations/implications – While the instrument shows promise, further research is needed to examine aspects of inter-rater reliability. Additional research is also needed to further examine relationships between it and other measures of research performance at the same level of analysis. While the usefulness and validity of this instrument at the "international level" has been examined, further research is needed to examine the relative validity and reliability of the instrument at the "institutional" and "national" levels.

Practical implications – The instrument provides a useful and cost-effective tool for use in the performance appraisal process of research scientists, and for use in focusing discussion on performance for developmental purposes. It is also useful as a research tool for the timely and cost-effective measurement of research performance at an institutional, national and international level.

Originality/value – The paper presents an original paper and pencil instrument for the appraisal of scientific research performance at an institutional, national, and international level.

Keywords Measurement, Research, Performance appraisal, Scientific management

Paper type Research paper

Introduction

Performance evaluation systems include elements such as appraisal measures and approaches for analyzing results which help in determining progress towards predefined objectives and assist in identifying areas for future improvement (Purbey *et al.*, 2007). As an element of performance evaluation systems, the use of multi-source feedback (including self-appraisals of performance) in the USA and the UK is growing (Fletcher and Baldry, 2000), while expanding the use of such systems in developing economies is also under consideration (Costigan *et al.*, 2005). Despite this apparent increase in the use of multi-source feedback there is some reluctance in the use of self evaluation of ability which has evolved from historical beliefs that individuals may

International Journal of Productivity and Performance Management
Vol. 58 No. 7, 2009
pp. 632-644

1741-0401
DOI 10.1108/17410400910989458

lack the objectivity and reliability needed to provide information on their own ability, and where self-appraisals are provided, they may be subject to considerable error due to a respondents desire for self-enhancement (DeNisi and Shaw, 1977; Levine *et al.*, 1977). More recent studies across various cultural groups still report quite poor levels of agreement between self and peers or supervisors ratings of job performance (Costigan *et al.*, 2005; Jaramillo *et al.*, 2005), which serves to reinforce such views. Additionally examination of the correlation between self-appraisals of performance and more "objective" measures of performance have also shown poor correlations (Pransky *et al.*, 2006).

While results are generally consistent in their findings of low to moderate correlations between self and others ratings it is important to recognise some of the limitation of many of these studies. The issue of restricted range is an important one. As measure variance increases so to will the correlations between the measures increase. This suggests the need for representative or random sampling in studies examining the relationship between self and other ratings. Unfortunately such selection techniques are infrequently used (Mabe and West, 1982). Other consideration such as the reliability of measures used, personal factors of the participants, such as intelligence levels, achievement status, and locus of control have also been found to influence the perceived accuracy of self-appraisals (Mabe and West, 1982). Additional factors such as lack of public information, the identification verses anonymity of the rater, poor self esteem, and the degree of likelihood that a favourable self-report will lead to personal gain, are also thought to influence the likelihood of self-enhancements, and consequently contribute to low inter-rater reliabilities (Schlenker, 1980).

Despite these concerns the uses of self-appraisal systems continue, and the value of including self-appraisal data as part the larger performance appraisal system is encouraged (Abu-Doleh and Weir, 2007). In addition to the evaluative purposes to which self-appraisals of work performance are put there is a growing secondary use as developmental tools. Increasingly researchers are highlighting the importance of incorporating self-appraisal data into the performance management systems employed in organisations.

Performance evaluation systems are a fundamental component of the application of human resource processes (Smith *et al.*, 1996). Campbell and Lee (1988) argue that utilising self-appraisals as developmental tools will improve job performance. Indeed the use of "360-degree" feedback from peers, superiors and subordinates can act as crucial contributors to an individuals diagnosis of their developmental needs (Parry and Sinha, 2005; Van Der Heijden and Nijhof, 2004). While the process of self assessment in and of itself may not lead to direct performance improvement (Andersen *et al.*, 2004), the success of employee development is based on effective needs assessment, and effective use of self-appraisal systems contributes significantly to that assessment (Alimo-Metcalfe and Lawler, 2001; Parry and Sinha, 2005))

Evaluating research performance

The measurement and evaluation of scientific productivity is a complex and sensitive area. How we define performance, evaluate quality over quantity, and quantify the potential value of scientific outputs serves to multiply the complexity of the issue. Essentially the measurement of scientific productivity is a measurement of knowledge

generation. Demarest (1997) highlighted the complexity of the process of generating knowledge and presented a model of knowledge management. According to the model, the knowledge flow consists of distinct steps such as knowledge construction, embodiment, dissemination, and use. Knowledge consists of both tacit and explicit elements. Scientists and innovators convert this tacit knowledge into explicit knowledge through the application of the scientific research process which helps in formalizing the body of knowledge (Nonaka and Takeuchi, 1995).

Various issues and difficulties have been highlighted in the literature with regard to measuring the effectiveness of knowledge generation or research productivity (Miller, 1982; Pappas and Remer, 1987; Vinkler, 1998; Werner and Souder, 1997). Such issues include absence of a single measure that can capture all dimensions of the scientific research process, the need to clearly specify measurement objectives, the difficulty in quantifying the research output, and sufficient detailed consideration for appropriate criteria, indicators, and data to be included in the measurement process.

In many cases knowledge generation cannot be measured directly; rather it is assessed indirectly by examining research outputs such as books, journal articles, reviews, patents, prototypes, etc. The complexity of such a task is illustrated by the detailed measurement of scientific productivity employed by the Unesco study, which initially identified 56 potential measures of research productivity (Andrews, 1979).

Within the academic domain common research performance measures include publication counts, citation indexing, complex bibliometric mapping ((Noyons, 2003; Van Raan, 1996), data envelopment analysis (Jyoti *et al.*, 2008), and multi-objective measurement (Nagpaul and Santanu, 2003). In addition to these approaches many individual countries employ government-initiated performance based measures to inform funding decisions. New Zealand's Performance Based Research Fund (PBRF) and the UK's Research Assessment Exercise (RAE) are examples of such measures. The logic behind such government initiatives is that the measurement of performance is a fundamental step both in the reward and the improvement and development of performance. While the performance measures mentioned previously offer useful "outwardly objective" measures of productivity they are, in most instances, costly, contentious, time-consuming, and lack flexibility beyond the express purposes for which they are designed.

While quantifiably "objective" measures of work performance are often perceived to represent the ideal in performance measurement, the complex nature of work in the modern knowledge economy, may result in such measures being elusive or indefinable (Tangen, 2003). Indeed disagreement within the scientific community on the validity of various measure of research performance is evident in the number of measures in existence.

The current study presents a measure of research performance based on scientists' self-appraisal of their research profile and performance. The development of this instrument originated from a need for a measure of research performance that would allow for the comparison of scientific performance between individuals and across scientific disciplines, while also maintaining the confidentiality of the scientists participating in the study. Additionally the measure of research performance was needed to be non-contentious to the participants, easy to administer and score, would not be resource intensive, and would be reliable and acceptably accurate.

Methodology

Measurement construction

The first stage of measurement construction involved the interviewing of research scientists to determine their perception of how research productivity and performance could be assessed; meeting the criteria outlined above.

A series of unstructured interviews were scheduled with several research active scientists ($n = 13$) from one British and one Irish University. The scientists themselves were from various disciplinary backgrounds (two physicists, four chemists, two mathematician, three biotechnologists, and two social scientists). A series of unstructured interviews were held with the general theme of the interview being "the measurement of research productivity". Notes were taken throughout each interview for later analysis. In some cases participants were contacted following their initial interview to discuss their comments in more detail and for feedback on the advances and variations to the measurement instrument being developed.

An analysis of notes taken during the interviews yielded two consistent findings. First it was generally reported by the participating scientists that no existing measure of research performance was viewed as universally satisfactory or acceptable. Second, given the continuing evaluation of a scientists research performance, that takes place throughout their career, at organisational, national and international levels; individual scientists believe they have a reasonably accurate perception of their "place" within the scientific community. While this first finding confirmed the need for additional research and development of measures of scientific performance the second finding also suggested that there could be some promise in a self-appraisal approach to the evaluation of scientific research performance.

A number of paper and pencil structures and formats of the measurement instrument were presented to the participating scientists. The results of feedback and comments from participants led to the refinement of the instrument known as the "Self-appraisal of research profile and performance" (SARPP) presented in the Appendix.

The instrument is based on the premise that in the process of a scientific career, a research scientist constructs a self-perceived research profile, which is a function of the quality and quantity of their work, and how it is received at various levels within the scientific community. During this ongoing process of evaluation and environmental feedback, scientists gain a clear understanding of how their profile reflects on them as a researcher and their "place" within the larger scientific community. In this context of self-appraisal, it is assumed that reasonably acceptable and standard criteria will be used as a benchmark with regard to quality journal lists and publication counts that are relevant within each scientific discipline. The instrument itself provides some descriptive guidelines for users to aid them in their self-appraisal. These guidelines, with reference to publication rates etc., are sufficiently vague to allow for variation in performance norms across different scientific disciplines.

The basic structure of the SARPP contains three sections, each with a Likert type five-point scale. The three sections relate to the respondents institutional profile, national profile, and international profile. The institutional profile relates to the respondents belief of where their "place" is within the research hierarchy of their institution. The national profile relates to the respondents sense of "place" at a national level, while the international profile relates to the scientists perception of their "place" within their research domain at an international level.

As the scientific community is essentially an international community, it is this international level that may be of primary significance, as this more clearly allows for the comparison of scientists across disciplines. For example we might imagine a situation whereby an average researcher working in a very narrow or novel field of research may quite reasonably assess themselves very highly at a institutional and national level of performance, simply due to the scarcity of other researchers in their field at the institutional and national levels. Or perhaps we could envision a situation where a researcher working in a predominantly teaching focused organisation might rate very highly at the institutional level because of the lack of research conducted by their institutional peers. However such researchers, in both examples would likely rate themselves much lower on the international level (given an accurate self-appraisal).

While the ultimate level of research appraisal is at the international level, the decision to include the other levels of performance (institutional and national) was made following discussion with the scientists participating in the development of the instrument. It was felt that the inclusion of the institutional and national levels would result in a more practicably versatile and useful instrument. Participating scientists also suggested that by including the first two levels, respondents' attentions would be more focused on the issue of "performance" which would lead to more reflective analysis of their position as researchers, and ultimately a more reliable measure.

Analysis of the SARPP

The examination of the validity and reliability of the SARPP required it to be completed by a sufficiently large number of research scientists with varying levels of ability. To ensure that research scientists of varying ability took part in the study a method for identifying scientists of varying ability was required. This was done by utilising a stratified random sample of scientists from biological and chemical science departments in UK university departments using the UK governments Research Assessment Exercise (RAE) as a grouping variable.

Selecting scientists of varying ability. The REA incorporates a wide range of evaluation criteria in its estimation of research performance, including such factors as research outputs, students and studentships, policies, evidence of esteem, and external research income. REA scores allow for the categorisation of researchers at the group level, within their departments from 5* to 1. As the aggregate score for a department is based on the submission of each research active member it is not unreasonable to equate the highest-ranking departments with the highest-quality research staff.

In addition to its usefulness as a "performance" measure the REA scores also facilitated stratification across performance levels during the sample selection phase of the study. In practical terms the REA measure allows for the identification of groups of researchers at various levels of performance, within specific research domains, across a large number of institutions. However this "group" labelling of research performance does not take into account individual differences, variations in performance levels among researchers in the same departments that may result from factors such as experience, age, gender, personality, qualification, length of service, or personal circumstance. Consequently the RAE provides a guide for the expected levels of performance of individual researchers while not offering a direct correlate of individual performance.

Validity and reliability of the SARPP. The issues of face validity and content validity were addressed during the construction of the instrument. The close collaboration and

repeated feedback from the original 13 research scientists who took part in the initial development stage of the instrument ensures some element of face and content validity for the likely participants which the instrument is intended for.

In addition to this, 270 research scientists (85.5 per cent male) who took part in their departments RAE completed the SARPP. Each participant was presented with a printout of the SARPP to be completed in their own time and returned to the investigator. Participants were also provided with a questionnaire to measure basic biographic data and work history. This data was then analysed to examine criterion validity: comparing the SARPP scores against REA scores, and representative reliability: examining known differences across groups.

The REA scores (5* to 1) for research departments are based on the aggregation of individual submissions across a department. Consequently one would expect that even allowing for individual differences and variation the aggregation of SARPP scores for participants who took part in the REA would correspond to the REA scores for that department.

This examination of criterion validity was conducted by comparing the mean scores for participants from each department against the REA scores for that department. It is important to note that only those participants who identified themselves as having their work included in their departments REA submission were included in this analysis. As the "International" level was the most senior and important evaluation level on the SARPP, participant's scores to this element of the SARPP were compared with the REA scores (see Table I).

It is important to note that while in theory the REA departmental scores range from 5* to 1, in practice relatively few departments are ranked at level one and two. It is probable that departments ranking at this level would not participate in the REA exercise, as the expenditure of time and resources required for submission would not be justified against the very low levels of research funding they would be awarded as a result. In order to assess the statistical significance of the relationship between mean SARPP and departmental RAE scores a Spearman's rho correlation was carried out.

As can be seen from Table II the relationship is significant at the 0.001 level, offering evidence in support of the criterion validity of the SARPP and the expected relationship between SARPP scores and REA scores. While the strength of the relationship is low to moderate it is consistent with other studies such as Jaramillo *et al.* (2005) which have examined the correlation between self-report and objective measures of performance.

In addition to this empirical measure of validity an additional measure of representative reliability was examined. The concept of representative reliability is based on the ability of an instrument to identify known differences across groups.

Departmental REA scores	Scientists mean scores for "international level" on SARPP	Number of participants in each group	Standard deviation
3b	2.6316	19	0.9551
3a	3.1176	68	0.9228
4	3.3000	70	1.0265
5	3.5614	57	0.9640
5*	3.8393	56	1.0579

Table I.
Participants mean scores for "international level" on SARPP across REA departmental scores

Ryan (2003) reports on a number of relationships shown to exist between demographic variables and scientific performance. One such, consistently identified relationship, is between scientific performance and age, with performance increasing as age increase and then dropping off slightly in the 60's and later years. The ability to identify this relationship through the analysis of SARPP data from the current study and a biographical measure of participant age would lend support to the representative reliability of the SARPP.

Since the ranges under which participants in the current study categorised themselves end at "60 + " it was not expected that the typically identified decline in performance would be present in the graphical illustration of age versus research performance. Consequently a steady increase of "international level" SARPP scores as age increases to the maximum "60 + " category was anticipated.

Figure 1 illustrates the increase in "international level" SARPP responses with age, as predicted offering some representative reliability to the SARPP. A Spearman correlation reports an $r = 0.342$, $p < 0.01$.This finding adds further evidence in support of the relative reliability of the SARPP.

Conclusion

The current study presents a measure of research performance based on a scientist's self-appraisal of their research profile and performance. The development of this instrument originated from a need for a measure of research performance that would

Table II.
Spearman's rho correlations for participants' mean scores for "international level" on SARPP and REA departmental scores

		REA	IntSARPP
REA	Spearman rho correlation	1	0.318
	Sig. (one-tailed)		0.001
	n	5	5

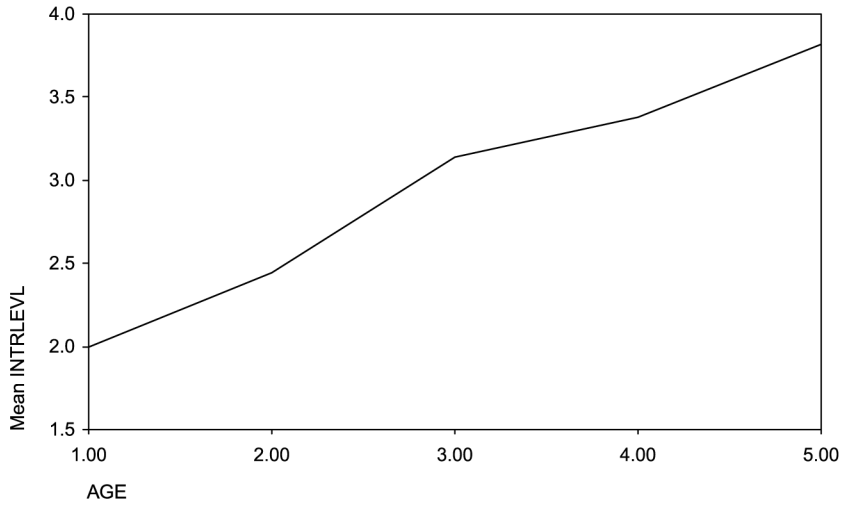

Figure 1.
Graph of age versus participant responses to "international level" on the SARPP

allow for the comparison of individual scientific performance between individuals and across scientific disciplines, and as such it met these requirements. Initial findings, though limited, offer some evidence of the SARPP as a valid self-appraisal tool.

The instrument is reported to have reasonable face validity and has been found to exhibit moderate criterion validity and representative reliability within the limits of its examination. Results are in line with reported examinations of the relationship between self-report measures of performance and objective performance criteria (Jaramillo *et al.*, 2005).

Participant feedback on the use of the instrument has been positive, with possible future uses including professional development planning for research scientists, organisational development for research institutions, as well its use as a convenient measure of individual research performance for non-critical research purposes, as was originally intended. It is important to recognise that depending on whether they are used for the purposes of performance evaluation, development or purely for research purposes error and bias may be more or less likely in self-report evaluations of performance feedback (London and Wohlers, 1991). As identified in the literature review, the success of employee development is based on effective needs assessment, and effective use of self-appraisal systems contribute significantly to that assessment (Alimo-Metcalfe and Lawler, 2001; Parry and Sinha, 2005). Results from the current study suggest that the SARPP could be usefully employed as a tool in the collection of data to support needs assessment and personnel development. Even allowing for possible bias and error in the self-appraisal process its use as a basis for discussion on performance issues still highlights its value (Inderrieden *et al.*, 2004), and the use of such measures in facilitating performance feedback and assisting in the establishment of future goals that can result in performance improvements (Stansfield and Longenecker, 2006).

It should also be noted that many of the intervening variables that are thought to moderate the relative accuracy of self-report measures of performance are of particular interest when considering a sample of research scientists. For example, relative accuracy in assessing ability has been linked with experience in the evaluation of that ability (Hodgson and Cramer, 1977). Evaluation of research in the form of the peer review process is a common occurrence for research scientists and is likely to lead to an improved ability to self-appraise performance. Similarly Mabe and West (1982) suggest that factors such as intelligence, which is expected to be high among a population of predominantly PhD qualified participants, should increase the likelihood of accurate self-appraisals.

Weakness of the SARPP should not be overlooked. While the instrument may serve as a focus for reflection and self-evaluation it must be recognised that ratings on their own do not communicate sufficiently detailed information to guide performance improvements. In order for development to be likely, participants in the evaluation need to recognise specific shortcomings or weaknesses and where changes need to occur (Van Der Heijden and Nijhof, 2004). With this in mind the SARPP may be useful as one aspect of a self-appraisal system that more completely looks at scientists performance across a number of important work areas and responsibilities. Thus allowing for greater flexibility in accounting for variables relating to individual differences, such as personal circumstances, differing research styles, differing research methods and strategies within and between disciplines, and a disciplines dependence on funding (Wood, 1990).

The correlation of aggregated SARPP scores with RAE departmental scores in this study is also problematic as it does not allow for the examination of relationships at the level of the individual for which the instrument is most likely to be used. Further research is encouraged, to examine individual responses to the SARPP and objective performance criteria, as well as the examination of paired self and supervisor ratings, and self and peer ratings on the SARPP. Importantly this will allow for comparisons of self and others ratings at the institutional, national and international levels; only international level ratings were able to be examined in any meaningful way in the current study.

The cultural context in which the current data is collected is also important. The scientific community is generally viewed as an international community. However, the majority of participants in the current study have come from a single national group and further research is necessary to determine if scientists from differing nations vary in the relative accuracy of their self-appraisals.

It should also be noted that the original 13 participants whose communications and conversations were utilised in the initial development of the SARPP were all scientists. As such their thoughts and perspectives on issues of performance measurement in science can be seen to come from a distinctly "scientific" perspective. There are of course other stakeholders interested in the performance of scientists and science and this measure does not purport to measure their perspective on performance in science. This relatively narrow focus is both a strength and weakness of the proposed instrument in that it offers a measure uniquely focused on scientists' perception of performance criteria while ignoring the performance criteria that other stakeholders in the science domain might view as important. While the instrument is limited in scope and focus it is hoped that it can contribute in a useful way to performance appraisal systems for scientists. It is also hoped that this instrument will stimulate further interest in the development of additional measurement approaches to supplement a more complete appraisal system for research scientists.

References

Abu-Doleh, J. and Weir, D. (2007), "Dimensions of performance appraisal systems in Jordanian private and public organizations", *International Journal of Human Resource Management*, Vol. 18, pp. 75-84.

Alimo-Metcalfe, B. and Lawler, J. (2001), "Leadership development in UK companies at the beginning of the twenty-first century: lessons for the NHS?", *Journal of Management Medicine*, Vol. 15, pp. 387-404.

Andersen, H.V., Lawrie, G. and Savic, N. (2004), "Effective quality management through third-generation balanced scorecard", *International Journal of Productivity & Performance Management*, Vol. 53, pp. 634-45.

Andrews, F.M. (1979), *Scientific Productivity: The Effectiveness of Research Groups in Six Countries*, Cambridge University Press, Cambridge.

Campbell, D.J. and Lee, C. (1988), "Self-appraisal in performance evaluation: development versus evaluation", *Academy of Management Review*, Vol. 13, pp. 302-14.

Costigan, R.D., Insinga, R.C., Kranas, G., Ilter, S.S., Kureshov, V.A. and Berman, J.J. (2005), "A study of Russian, Polish, and US self- and peer raters: moderators of rating agreement", *International Journal of Commerce & Management*, Vol. 15, pp. 205-20.

Demarest, M. (1997), "Understanding knowledge management", *Long Range Planning*, Vol. 30, pp. 321-2, 374-84.

Denisi, A.S. and Shaw, J.B. (1977), "Investigation of the uses of self-reports of abilities", *Journal of Applied Psychology*, Vol. 62, pp. 641-4.

Fletcher, C. and Baldry, C. (2000), "A study of individual difference and self-awareness in the context of multi-source feedback", *Journal of Occupational and Organizational Psychology*, Vol. 73, pp. 303-19.

Hodgson, M.L. and Cramer, S.H. (1977), "The relationship between self-estimated and measured abilities in adolescents", *Measurement and Evaluation in Guidance*, Vol. 10, pp. 98-105.

Inderrieden, E.J., Allen, R.E. and Keaveny, T.J. (2004), "Managerial discretion in the use of self-ratings in an appraisal system: the antecedents and consequences", *Journal of Managerial Issues*, Vol. 16, pp. 460-82.

Jaramillo, F., Carrillat, F.A. and Locander, W.B. (2005), "A meta-analytic comparison of managerial ratings and self-evaluations", *Journal of Personal Selling & Sales Management*, Vol. 25, pp. 315-28.

Jyoti, T., Banwet, D.K. and Deshmukh, S.G. (2008), "Evaluating performance of national R&D organizations using integrated DEA-AHP technique", *International Journal of Productivity and Performance Management*, Vol. 57, pp. 370-88.

Levine, E.L., Flory, A. and Ash, R.A. (1977), "Self-assessment in personnel selection", *Journal of Applied Psychology*, Vol. 62, pp. 428-35.

London, M. and Wohlers, A.J. (1991), "Agreement between subordinate and self-ratings in upward feedback", *Personnel Psychology*, Vol. 44, pp. 375-90.

Mabe, P.A. and West, S.G. (1982), "Validity of self-evaluations of ability: a review and meta-analysis", *Journal of Applied Psychology*, Vol. 67, pp. 280-96.

Miller, R.B. (1982), "Measurement issues in R&D productivity", in Fusfeld, H.I. and Langlois, R.N. (Eds), *Understanding R&D Productivity*, Pergamon, New York, NY.

Nagpaul, P.S. and Santanu, R. (2003), "Constructing a multi-objective measure of research performance", *Scientometrics*, Vol. 56, pp. 383-402.

Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, New York, NY.

Noyons, E. (2003), "Evaluative bibliometrics as a science policy supportive tool", in Hurley, J. (Ed.), *Scientific Research Effectiveness: The Organisational Dimension*, Kluwer Academic Publishers, Dordrecht.

Pappas, R. and Remer, D.S. (1987), "Measuring R&D productivity", in Timpe, D. (Ed.), *Productivity: The Art and Science of Business Management*, Kend Publishers, New York, NY.

Parry, K.W. and Sinha, P.N. (2005), "Researching the trainability of transformational organizational leadership", *Human Resource Development International*, Vol. 8, pp. 165-83.

Pransky, G., Finkelstein, S., Berndt, E., Kyle, M., Mackell, J. and Tortorice, D. (2006), "Objective and self-report work performance measures: a comparative analysis", *International Journal of Productivity and Performance Management*, Vol. 55, pp. 390-9.

Purbey, S., Mukherjee, K. and Bhar, C. (2007), "Performance measurement system for healthcare processes", *International Journal of Productivity and Performance Management*, Vol. 56, pp. 241-51.

Ryan, J.C. (2003), "The contribution of cognitive and organisational psychology", in Hurley, J. (Ed.), *Scientific Research Effectiveness: The Organisational Dimension*, Kluwer Academic Publishers, Dordrecht.

Schlenker, B.R. (1980), *Impression Management: The Self-concept, Social Identity, and Interpersonal Relations*, Brooks/Cole, Monterey, CA.

Smith, B.N., Hornsby, J.S. and Shirmeyer, R. (1996), "Current trends in performance appraisal: an examination of managerial practice", *SAM Advanced Management Journal*, Vol. 61, pp. 10-15.

Stansfield, T.C. and Longenecker, C.O. (2006), "The effects of goal setting and feedback on manufacturing productivity: a field experiment", *International Journal of Productivity and Performance Management*, Vol. 55, pp. 346-58.

Tangen, S. (2003), "An overview of frequently used performance measures", *Work Study*, Vol. 52, pp. 347-54.

Van Der Heijden, B.I.J.M. and Nijhof, A.H.J. (2004), "The value of subjectivity: problems and prospects for 360-degree appraisal systems", *International Journal of Human Resource Management*, Vol. 15, pp. 493-511.

Van Raan, A.F.J. (1996), "Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises", *Scientometrics*, Vol. 36, pp. 397-420.

Vinkler, P. (1998), "General performance indexes calculated for research institutes of the Hungarian Academy of Sciences based on scientometric indicators", *Scientometrics*, Vol. 41, pp. 185-200.

Werner, B.M. and Souder, W.E. (1997), "Measuring R&D performance – state-of-the-art", *Research Technology Management*, Vol. 40, p. 34.

Wood, F. (1990), "Factors influencing research performance of university staff", *Higher Education*, Vol. 19, pp. 81-100.

Further reading

Mersman, J.L. and Donaldson, S.I. (2000), "Factors affecting the convergence of self-peer ratings on contextual and task performance", *Human Performance*, Vol. 13, pp. 299-322.

Appendix. Self-appraisal of research profile and performance

Research on self-appraisal suggests that research scientists working in the university setting have an accurate perception of their position within the scientific community.

This measurement instrument utilises this perception by allowing scientists to evaluate their position in the scientific community across three dimensions:

(1) Institutional;

(2) National; and

(3) International.

All responses given in the following sections will be held in the strictest confidence.

Research performance and profile at the institutional level

The institutional dimension refers to your standing as a researcher in your institution, with 5 being the very highest score and 1 being the very lowest score.

For example a score of 5 may refer to an individual who has a high publication rate in comparison to their peers, whose research is well regarded and respected by their colleagues, and is generally seen as one of the "shining lights" of the research faculty. Such individuals may have also won special research awards within their institution.

A score of 3 may represent an individual who has an average but consistent level of research productivity, and who contributes in a moderate way to the research profile of the institution in which they work.

A score of 1 may represent an individual who, for whatever reason, has not published work for several years, is not active on a current research project and does little to contribute to the research profile of their institution.

I would now ask you to rate yourself on the five point scale below, keeping in mind the examples given above.

Institutional profile and performance

5 4 3 2 1

Research performance and profile at the national level

The national dimension refers to your standing as a scientific researcher on a national level. Again with 5 being the highest score and 1 being the lowest.

For example a score of 5 may refer to an individual whose research would be well know and respected by his/her colleagues across the country. He/she would be easily recognised as the national expert in their area of specialisation. Colleagues working in the same research area may also consistently refer to his/her work. They may also have won national science or research awards, and large grants to fund their research.

A score of 3 may represent an individual who would be known nationally for their contributions to a specialist area, but may not be necessarily seen as a leader in that field. Their work may receive occasional citation from researchers in the same area, and they may be successful in attaining moderate funding for their research.

A score of 1 may represent an individual whose research is relatively unknown. They may produce a very small number of research publications that are not widely cited. They may also be generally unknown on a national level outside of a small group of colleagues from other institutions.

I would now ask you to rate yourself on the five-point scale below, keeping in mind the examples given above.

National profile and performance

5 4 3 2 1

Research performance and profile at the international level

The international dimension refers to your standing as a scientific researcher on an international level. Again with 5 being the highest score and 1 being the lowest.

For example a score of 5 may refer to an individual whose research would be well know and respected by his/her colleagues across the world. He/she would be easily recognised as the international expert in their area of specialisation. Colleagues working in the same research area would consistently refer to his/her work. They may also have won international science or research awards, or have been nominated or short listed for such awards.

A score of 3 may represent an individual who would be known internationally for their contributions to a specialist area, but may not be necessarily seen as the leader in that field. Their work would be regularly cited from researchers in the same area, and they would be successful in attaining funding for their research.

A score of 1 may represent an individual whose research is relatively unknown internationally. They may produce a very small number of research publications that are not widely cited. They may also be generally unknown on an international level outside of a small group of colleagues from other institutions.

I would now ask you to rate yourself on the five-point scale below, keeping in mind the examples given above.

International profile and performance

5 4 3 2 1

Thank you for the time and effort you have spent on completing this self-evaluation exercise. The responses you have given are an integral part of my research. I would once again like to ensure participants that all the information gathered will be held in the strictest confidence and no individuals will be identified by the study.

About the authors

James C. Ryan holds a PhD from Dublin City University Business School, and a BSc and MA from the National University of Ireland, University College Dublin. He is as assistant professor in the Department of Management, Marketing and Public Administration, University of Sharjah, UAE and former Senior Lecturer at Massey University, New Zealand. He is the recipient of a Government of Ireland Scholarship from the Irish Research Council for the Humanities and Social Sciences. His research areas are mainly in organisational behaviour and human resource management, with a focus on scientific productivity and the organisational environment in which knowledge workers operate. James C Ryan is the corresponding author and can be contacted at: jcryanphd@gmail.com

Syed Awais Tipu holds a DSc and MSc from the George Washington University. He is an assistant professor in the Department of Management, Marketing and Public Administration, University of Sharjah, UAE. His research areas are mainly in innovation management, organizational culture, leadership, and motivation.