

Implementation of a scientometric-based evaluation system for research units at a Romanian University

RADU SILAGHI-DUMITRESCU, SIMION SIMON, SORIN FILIPESCU, AUGUSTA PLOSCARIU
Babes-Bolyai University, 1 M. Kogalniceanu str, Cluj-Napoca 400028, Romania

The Babes-Bolyai University in Romania (BBU), with several thousand students and ~1500 faculty members, is one of the largest universities in south-eastern Europe, and one of the higher-ranked universities at national level in terms of scientific research, with a comprehensive profile that spans exact sciences, social sciences, economics, theology and others. A scientometric-based evaluation system was recently applied at this university for its research groups, taking into account ~50 items presumed to afford quantification of scientific output, and defining two separate sets of formulae for exact sciences vs the rest of the disciplines. The protocol and the results of this evaluation are discussed, highlighting the relative contributions of the components to the overall score.

Keywords: Scientometric, Romania, Babes-Bolyai, Research, Ranking, Evaluation

Institutional context

The Babes-Bolyai University at Cluj-Napoca Romania, has a comprehensive profile, with 21 Faculties spanning mathematics, physics, chemistry, biology, geology, geography, environmental sciences, history, philosophy, linguistics, theater, television, psychology, sociology, political sciences, European studies, business administration, economics, sports sciences, theology (of several different confessions), and others. These encompass ~1500 faculty, ~40000 students, and one of the longest traditions in Eastern Europe, doubled by a tradition for highest rankings on national level [1-3]. Within this diverse community, there is a clear pressure to define performance criteria for research, with as little discipline-bias as possible. Two opposing philosophies may be employed to this end: a qualitative approach relying on competent peer-review, versus an (apparently) impersonal ‘golden-number’ approach, where numerical criteria are proposed trying to encompass the diversity of types of outputs provided by a wide range of research fields. Societies recently emerged from totalitarian regimes, and/or where repeated sudden social changes have occurred, may be expected to have a natural tendency to distrust institutional approaches whose outcome depends on well-defined individuals (e.g., peer-review), and instead favor the less personalized numerical approach. Such distrust would be driven by prior experiences where, by the very nature of totalitarian regimes, institutional

decisions have had a tradition to not always be put in the most competent hands available. It is in this context that the Babes-Bolyai University has experimented a system of research evaluation for units/groups, based on comprehensive numerical criteria.

Institutional procedures

The system, based on ideas originating from a think-tank of the Romanian Academy, was devised by a commission of the Senate of the University, submitted for debate for 90 days to all the staff. Following the suggestions and comments formulated by individuals and by heads of Departments/Faculties, the Senate (consisting of ~120 representatives of all 21 Faculties as well as of the student body) approved the criteria almost unanimously. A separate committee (the Scientific Council of the University) was appointed to oversee the implementation of the criteria, aided by two technical staff from the University’s administrative offices. Research groups were then invited to fill in digital forms, produced in a general-purpose commercial software package. These forms were pre-completed with generally-available information, or with information already held by the University, in order to minimize the amount of information handled by the researchers. The forms (attached as Supporting Information) also embed pre-calculated formulas, so that the researchers did not need to

perform mathematical calculations manually: divisions by the number of authors, multiplications/divisions by set coefficients, weighted sums and other operations are all performed automatically; moreover, journal impact factors are introduced by dedicated staff, rather than by research groups. The forms also contained a free-format section, allowing the groups to cover issues not addressed by the numerical data; additionally, definition of new numerical criteria was also allowed, provided they were backed up by the Departments/Faculties and that their weight would not be larger than the average weight of the other criteria.

Each research unit was asked to choose one of three evaluation forms: ‘exact, engineering and biomedical sciences’ (Domain I) vs. ‘social and humanities,’ (Domain II) and “economic sciences” (Domain III).

Participation in the evaluation procedure was voluntary, both in terms of the research units as entities, and in terms of each member of each unit; written consent was collected. Each research unit submitted their evaluation form to the university’s Department for Management of the Scientific Research, completed with data not otherwise available to the administrative offices of the university, as detailed in the next section. These forms were then checked for consistency, and formulae were calculated, by the administrative specialized staff. Two sets of evaluators were then defined – one set for Domain I and one set for Domains II + III (only one research group from the economics domain expressed an interest in evaluation); the reviewers were proposed by the Deans of the Faculties, and approved by the Scientific Council. Each project was evaluated anonymously by three reviewers, randomly selected from within the evaluated domain, with limitations pertaining to standard potential conflicts of interest. The three evaluators were in contact with each other, and asked to come to a consensus in terms of the overall numerical score; they were further asked to recommend whether the research unit should receive an ‘accreditation certificate’ from the university, to certify that the respective research unit is deemed to be representative for the university.

The consensus scores and recommendations were then discussed by each of the two bodies of evaluators. The Scientific Council then merged the

data from the two Domains and submitted a final report to the Senate, which approved it almost with unanimity.

Criteria

For all three Domains, the evaluation form began by a simple text section describing administrative issues connected with the research unit (sections 1 and 2). Section 3 would then describe:

- the ‘mission of the research unit, its directions of research, development and innovation, and excellence results’ (section 3.1, with a limit of 2000 characters);
- brief general comments on the ways in which the results of the unit have been disseminated (section 3.2, maximum 1000 characters) – e.g., mentioning results in international classifications where available, members with unusually high numbers of citations, and others;
- comments on the financial situation of the research unit (section 3.3) ;
- general overview of the personnel (section 3.4), indicating formal degrees (professor, associate professor, assistant professor, research assistant, PhD student, technician, etc.) ;
- the total number of members of the research unit, N_c (section 3.5), to be used in formulae throughout the forms.

Sections 4-8 contain the numerical and scientometric criteria. It is noteworthy that the evaluators were also allowed to increase the total score of the research unit by at most 15% for unusually important arguments raised in Section 3, if other comments were considered necessary. Additionally, the reviewers were also allowed to increase the score by up to 10% for remarkable achievements (examples given were: papers cited for more than 100 times excluding self-citations, or cover images in Web-of-Science journals).

Tables 1 to 5 show the criteria for Sections 4-8 related to all three Domains. The total score within each Domain was calculated as a weighted average, with the following coefficients: 0.4 for Section 4, 0.2 for Section 5, 0.1 for Section 6, 0.1 for Section 7, and 0.2 for Section 8; alternatively, a simple, non-weighted total was also made available for comparison.

Table 1. Data and formulae for Section 4 – “Criteria of performance in scientific research for 2008-2012”.

No.	Criterion	n (number of items)	Points per item (Domain specified in parentheses)	Points per section ⁴
1.	Scientific book in a reputed publishing house abroad (author book, academic critical edition, critical edition of a text/manuscript with issues connected with interpretation of (hand)writing, translation from classical languages) ¹		$20 \times (N_{ic}/N_a)$ (I, II, III)	
2.	Monograph with a reputed publisher abroad ¹		$15 \times (N_{ic}/N_a)$ (I, II, III)	
3.	Article indexed in <i>Web of Science</i> (Thomson Reuters) ² , or ERIH, or recognized within the field of Humanities		$(10 + FI) \times (N_{ic}/N_a)^3$ (I, II, III)	
4.	Full paper in conference Proceedings, indexed in <i>Web of Science</i> (Thomson Reuters)		$(10 + FI) \times (N_{ic}/N_a)^3$ (I, II, III)	
5.	Full paper at an international scientific meeting, published by a reputed publishing house abroad, including electronic versions		$2 \times (N_{ic}/N_a)$ (I, II, III)	
6.	Chapter in a treatise, book or monograph, with a reputed publisher abroad ^{1,3}		$13 \times (N_{ic}/N_a) \times (N_p/N_{tp})$ (I, III) $15 \times N_{ic}/N_a \times N_p/N_{tp}$ (II)	
7.	Number of citations according to <i>Web of Science</i> (Thomson Reuters)		1/citation (I, II, III)	
8.	Sum of impact factors according to <i>Web of Science</i> (Thomson Reuters) ²	-	$\Sigma (FI \times (N_{ic}/N_a))$ (I, II, III)	
9.	Book with a reputed publisher within Romania ⁵		$7 \times (N_{ic}/N_a)$ (I, II, III)	
10.	Article in a journal indexed by the Romanian National Council for Scientific Research as B+, or indexed in an international database (these typically include journals indexed in reputed international databases of more specialized scope than <i>Web of Science</i> ; in Domain III the databases RepeC, DOAJ, EPSCO, IndexCopernicus, Scopus, ScienceDirect, Wiley-Blackwell Full Collection, SpringerLink, Emerald, Inderscience by exception receive 5x points)		$1 \times (N_{ic}/N_a)$ (I, III) $2 \times (N_{ic}/N_a)$ (II) $5 \times (N_{ic}/N_a)$ (III-exceptions)	
11.	Invited/plenary/keynote conference at an international scientific meeting		10 (I, II, III)	
12.	Invited/plenary/keynote conference at a national scientific meeting		5 (I, II, III)	
13.	Oral presentation at an international scientific meeting		$5 \times (N_{ic}/N_a)$ (I, II, III)	
14.	Oral presentation at a national scientific meeting		$2 \times (N_{ic}/N_a)$ (I, II, III)	
15.	Participation in a national program (e.g., for archeology)		$20 \times N_{ic}/N_a$ (II)	
16.	Participation to a national or international exposition, with heritage exhibits		$15 \times N_{ic}/N_a$ (II)	
17.	Management and conservation of a cultural and/or scientific archive/fund of national and/or international interest		$20 \times N_{ic}/N_a$ (II)	
18.	A reference book (collective authorship) in a reputed foreign publishing house (dictionary, lexicon, encyclopedia, atlas, index of testimonials, and others) ²		$20 \times N_{ic}/N_a$ (II)	
19.	Translation of a scientific book, published at the Romanian Academy's Publishing House, or in a reputed national publishing house, with an introductory specialist study ¹		$15 \times N_{ic}/N_a$ (II)	
20.	Review published in a specialized journal abroad, indexed ISI, ERIH, or recognized within the field of Humanities		$2 \times N_{ic}/N_a$ (II)	
21.	Review published in a specialized national journal indexed by the Romanian National Council for Scientific Research as B+ or B, or indexed in an international database (these typically include journals indexed in international databases of more specialized scope than <i>Web of Science</i>)		$1 \times N_{ic}/N_a$ (II)	
22.	Scientific study or specialized report connected to a research program, published on the internet or on electronic support.		$2 \times N_{ic}/N_a$ (II)	
23.	A specialized journal published by the research unit		20 (II)	

No.	Criterion	n (number of items)	Points per item (Domain specified in parentheses)	Points per section ⁴
24.	Article published on the site of the European Commission, ILO, OCDE, Eurofound		$5 \times (N_{ic}/N_a)$ (III)	Σ
25.	Number of citations in journals indexed in other international databases (RePeC, DOAJ, EPSCO, IndexCopernicus, Scopus, ScienceDirect, Wiley-Blackwell Full Collection, SpringerLink, Emerald, Inderscience, etc.)		0.2/citation (III)	

¹ These are defined as: Academic Press, Appleton & Lange, Birkhauser, Blackwell, Cambridge University Press, CRC Press, Elsevier, Garland Publishing, Kluwer Academic Publishers, McGraw-Hill, Mosby, Nova Science Publishers, Oxford University Press, QMP, Springer Verlag, Thieme, Willey-Liss, Williams and Wilkins, World Scientific Publishing, or other international publishing houses of similar reputation. If in several volumes, each volume is treated as a separate item and scored as such.

² For each article, the impact factor (IF, taken from Thomson Reuters' *Web of Knowledge, Journal Citation Report* edition for the year when the article was published) is divided by the total number of authors (N_a) and multiplied by the number of authors affiliated at that time with the evaluated research unit (N_{ic}). For Domain II, only foreign journals are counted under this criterion.

³ n = number of items (programs, books, chapters, articles, citations, etc.); N_p = number of pages per chapter; N_{tp} = number of pages per volume.

⁴ Simple sum of all individual items for the respective criterion.

⁵ Scientific books authored by the researcher, indexed in the National Library of Romania.

Table 2. Data and formulae for Section 5 – “The ability to attract research funds, 2008-2012”; for joint grants only the exact amounts directly overseen/spent by the members of the research unit are considered).

No.	Criteria	n	Points per item (domain in parentheses)	Points given
1.	Grants won by the research unit from international funding bodies	5 000 - 10 000 EUR	2 (I, II) 10 (III)	
		10 001 - 50 000 EUR	4 (I, II) 15 (III)	
		50 001 - 200 000 EUR	6 (I, II) 20 (III)	
		200 001 - 1 000 000 EUR	8 (I, II) 25 (III)	
		> 1 000 000 EUR	10 (I, II) 30 (III)	
2.	Grants won by the research unit from national funding bodies	< 10 000 RON	1 (I, II) 5 (III)	
		10 001 - 100 000 RON	2 (I, II) 10 (III)	
		100 001 - 500 000 RON	3 (I, II) 12 (III)	
		> 500 000 RON	4 (I, II) 15 (III)	
3.	Contracts for external funding obtained by the research unit from international or national organizations	< 5 000 RON	0,5 (I, II) 10 (III)	
		5 001 - 10 000 RON	1 (I, II) 12 (III)	
		10 001 - 100 000 RON	2 (I, II) 15 (III)	
		< 100 000 RON	3 (I, II) 20 (III)	
4.	International scientific meeting (congress, conference, symposium) or summer school organized by the research unit.		10 (I, II) 20 (III)	
5.	National scientific meeting (congress, conference, symposium) or summer school organized by the research unit.		5 (I, II) 10 (III)	

Table 3. Data and formulae for Section 6 – “The ability to develop services, technologies, products, 2008-2012”.

No.	Criteria	n	Points per item (domain in parentheses)	Points given
1.	Registered request at a patent office	International office	5 (I, III)	
		National office	2 (I, III)	
2.	Patent granted	international	10 (I, III)	
		national	5 (I, III)	
3.	Patent applied	internationally	20 (I, III)	
		nationally	10 (I, III)	
4.	Patent cited in <i>Web of Science</i> (Thomson Reuters)		5 (I, III)	
5.	Products and technologies resulted from internally-developed research, innovation and development (products sold, with payments received by the research unit)		10 (I) 20 (III)	
6.	Development and accreditation of a research laboratory		10 (I)	
7.	Impact studies ordered by external beneficiaries		5 (I, III), 10 (II)	
8.	Specialized report prepared for competent Government bodies, or for non-governmental organizations		20 (II)	
9.	Consultancy/research on specialized topics/issues, requested by official institutions and given in the form of written reports		10 (II) 20 (III)	

Table 4. Data and formulae for Section 7 – “The ability to provide higher education and training to young researchers - 2008-2012”.

No.	Criteria	n	Points per item (domains in parentheses)	Points given
1.	PhD advisors working within the research unit		20 (I, II, III)	
2.	PhD students working within the research unit		10 (I, II, III)	
3.	Post-docs working within the research unit		20 (I, II)	
4.	Researcher who has defended their PhD while being a member of the research unit		20 (I, II, III) 10 (III)	
5.	Members in advisory committees for PhD students		5 (I, II, III)	
6.	Ratio of Ntd (people who have defended their PhD less than 10 years ago) / total members of the research unit Nc		$100 \times \text{Ntd} / \text{Nc}$ (I, II, III)	

Table 5. Data and formulae for Section 8 – “Scientific prestige” (no time limit, unlike the rest of the criteria), in Domains I and II.

No.	Criteria	n	Points per item	Points given
1.	Member in editorial boards of national/international journals indexed by <i>Web of Science</i> - Thomson Reuters or by other international databases, or in the editorial board of reputed publishing houses		30 (I, III)	
2.	Member in the management of an international professional association		20 (I, II, III)	
3.	Member of the Romanian Academy, or of other national academies		50 (I, II, III)	
4.	Researchers with Hirsch index > 8		30 (I)	
5.	Honorary member (including fellow, senior) of a national/international scientific society		20 (I, II, III)	
6.	Prize received from the Romanian Academy		30 (I, II, III)	

No.	Criteria	n	Points per item	Points given
7.	Prizes from international scientific societies, when received based on a selection process		20 (I, II) 30 (III)	
8.	Guest editor for a national/international journal indexed in <i>Web of Science</i> - Thomson Reuters or in other international database		15 (I, II, III)	
9.	Reviewer for a scientific national/international journal indexed in <i>Web of Science</i> - Thomson Reuters or in other international databases		5 (I, II, III)	
10.	Prizes from national scientific societies, when received based on a selection process		5 (I, II) 10 (III)	
11.	Members in defense committees for PhD theses		5 (I, II, III)	
12.	Member in editorial boards of national/international journals		30 (II)	
13.	Member in editorial boards of journals edited by the Romanian Academy, or other national journals of recognized reputation in their fields		20 (II)	
14.	<i>Doctor honoris causa</i>		20 (II)	
15.	Member in international professional association		20 (II)	
16.	Researchers with Hirsch index > 4		20 (II, III)	
17.	Member in the management of a national professional association		10 (II)	
18.	Prizes and distinctions from government institutions/authorities		30 (III)	
19.	Diplomas of excellence issued by scientific national and international organizations		20 (III)	

Analysis of results

Upon examining the results of the evaluation, the Scientific Council as well as the two bodies of evaluators concluded that the criteria should be improved before any measures are taken based on numerical classifications derived from the present evaluation. The numerical data is summarized in Figures 1-4 and in Tables 6 and 7.

700 of the ~1500 permanent faculty and researchers of the University were present within the research units evaluated by the procedure described here. A significant part (but not all) of the remaining faculty is dedicated almost exclusively to teaching activities. Only ~60 of these 700 are full-time researchers, and they are in majority members of one single research unit.

Fifty-six research units were evaluated, with scores between 100 and 4000 points, with an average score of ~700. The highest scores came from structures defined as Institutes.

The weighted-sum score, described above, matches very well an alternative score calculated as a non-weighted sum - $R^2=0.98$ cf. Figure 1.

Section 4 correlates very well with the total score ($R^2=0.9711$). By contrast, the very small values registered for Sections 5 and 6 suggest that in their

present form they were superfluous as part of the total/aggregate score.

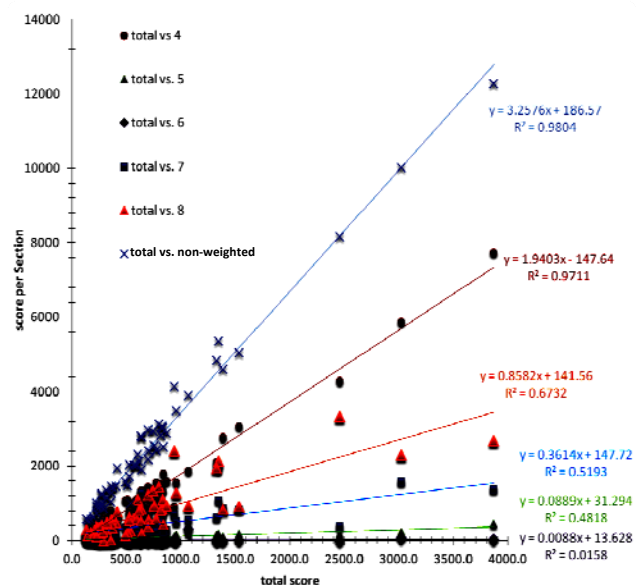


Fig. 1. Correlations between total scores and the scores obtained at each of the Section 4-8, for the 56 research units evaluated.

As shown in Fig. 2, the total score also correlates to some extent with the size of the research unit.

Thus, the total number of members of the research unit (N_c) has an R^2 of 0.64; on the other hand, the correlation disappears when one leaves out the temporary employees (students and postdocs) - as R^2 becomes 0.12.

The average score per researcher is at ~ 70 , with individual values ranging from 10 to 170. There is a

reasonable correlation between the total score and the number of Thomson-Reuters (ISI) papers, $R^2 = 0.80$; this correlation is more obvious for the highest scores. In this respect, there is an average of 0.7 ISI papers per person per year for the 700 participants.

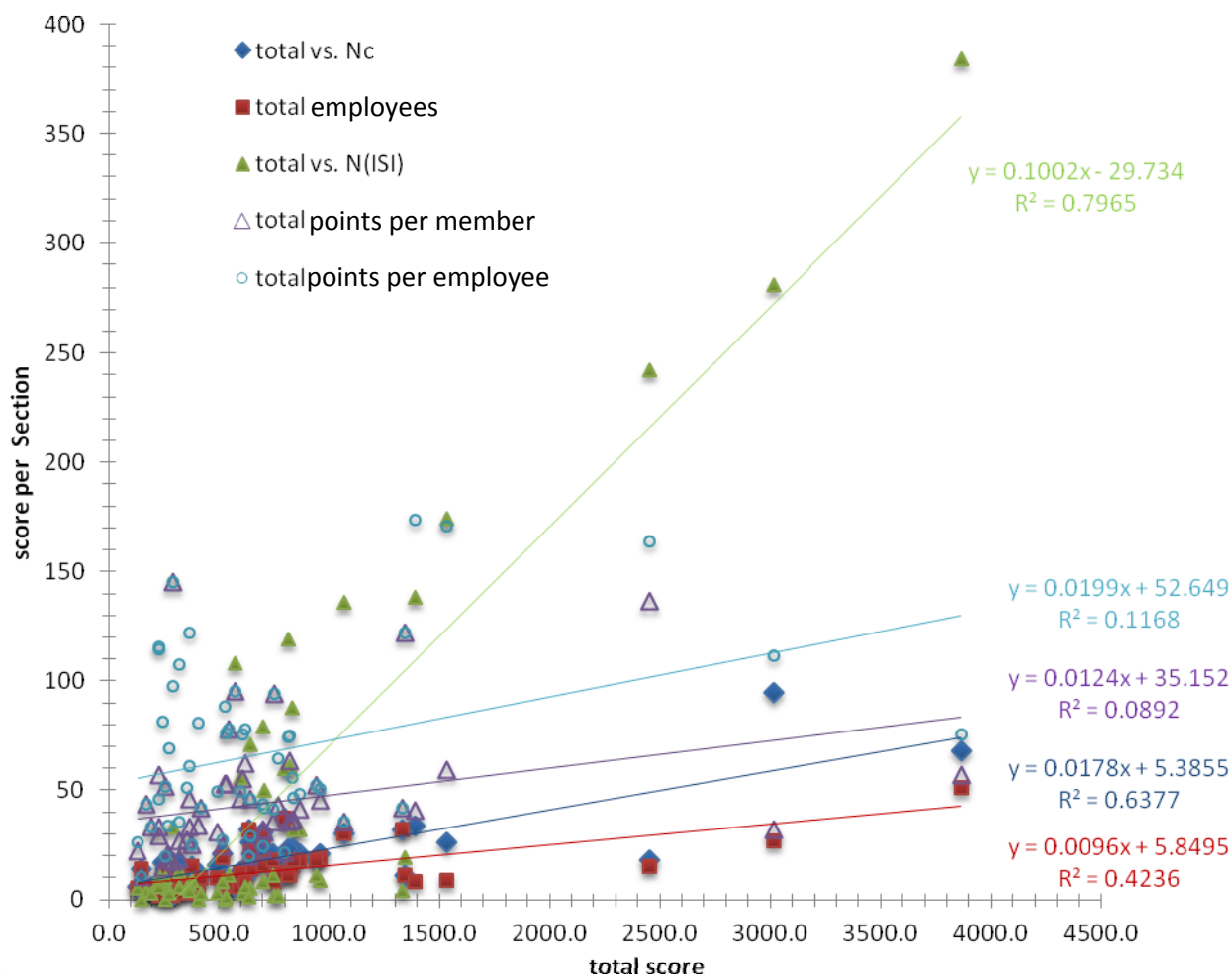


Fig. 2. Correlations between total scores and the number of members of the research unit, the number of permanent employees, and the number of papers indexed in Thomson-Reuters (ISI).

Fig. 3 shows the same statistics as in Figs. 1 and 2, now limited only to the research units within Domain I; the Domain II data is summarized in Fig. 4. For Domain I, the correlation between Section 4 and the total score is excellent ($R^2 = 0.99$), suggesting that Sections 5, 6, 7 and 8 may have been redundant – or that their weight should be increased. In fact, Sections 5 and 7, while yielding numerical values significantly smaller than Section 4, do show a reasonable correlation with the overall score – and implicitly with Section 4 as well. On the other hand, Section 6 shows zero correlation with the total score. This, together with the very small values registered by all research units at this Section, suggests a hiatus

between the general research efforts within the Babes-Bolyai University, and the activities covered by Section 6 – namely, applied research. By contrast, the Domain I scores correlate very well with the total number of ISI papers, but only poorly with the number of members.

For Domains II+III, the correlation between the total score and Section 4 is significantly smaller, with R^2 dropping down to 0.85 and now equaled in weight by Section 8; the latter shows similar degrees of correlation with the total scores in Domain I vs. II+III ($R^2 \sim 0.7$). The number of members, or the number of ISI papers, show negligible correlation with the total score.

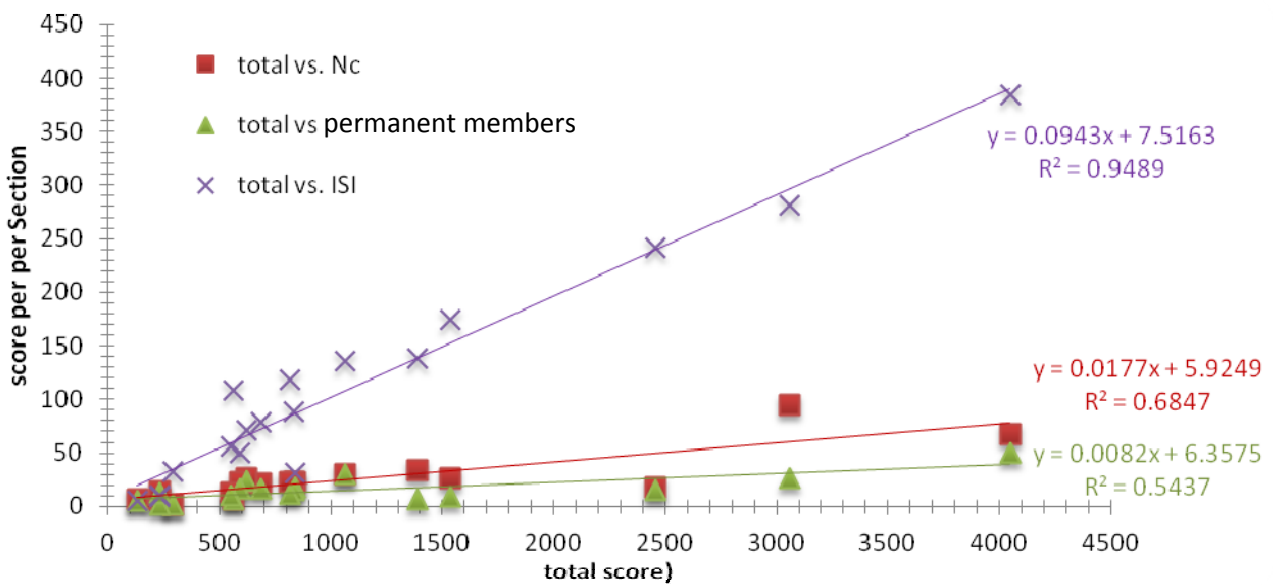
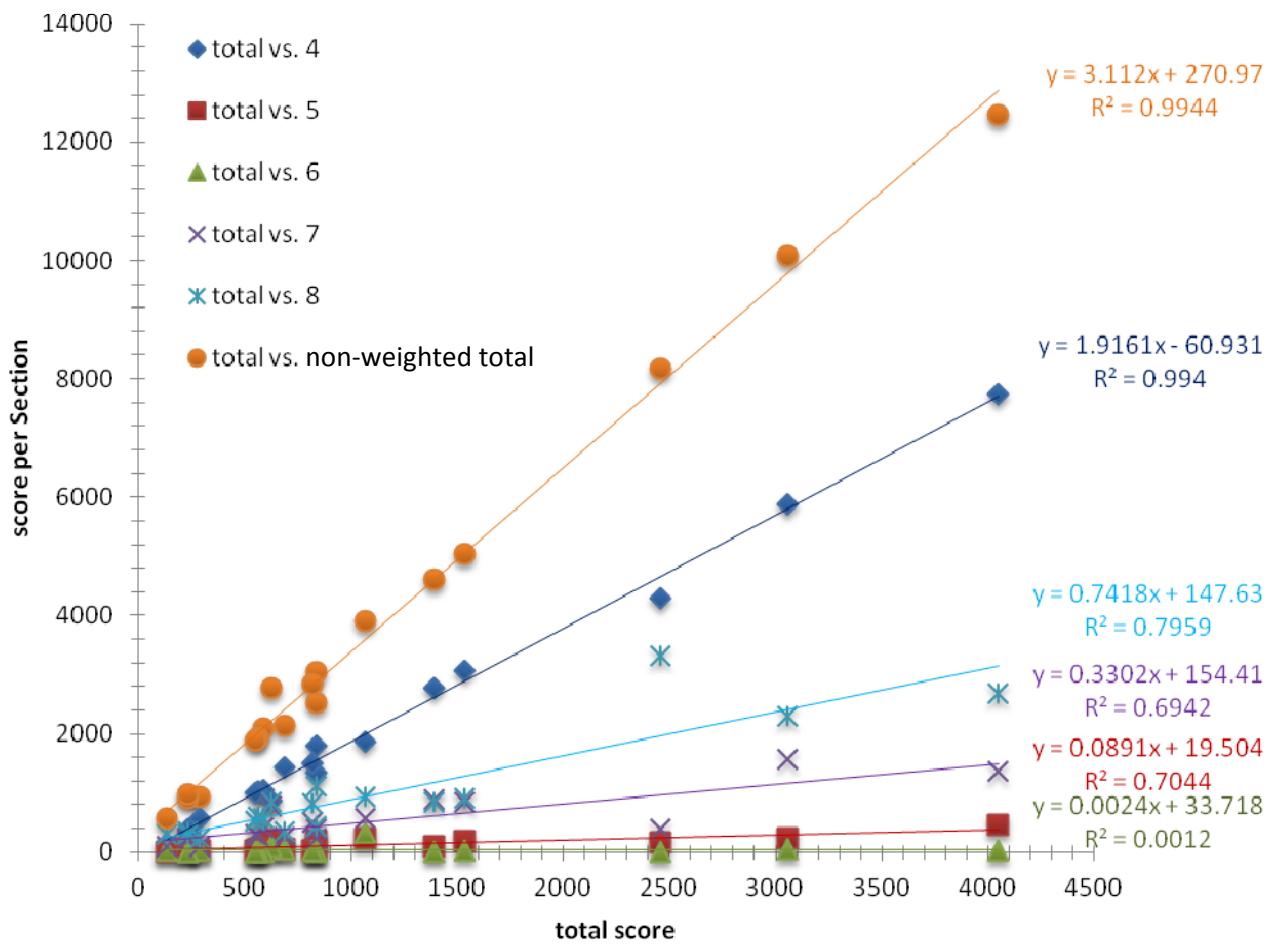


Fig. 3. Statistics for Domain I.

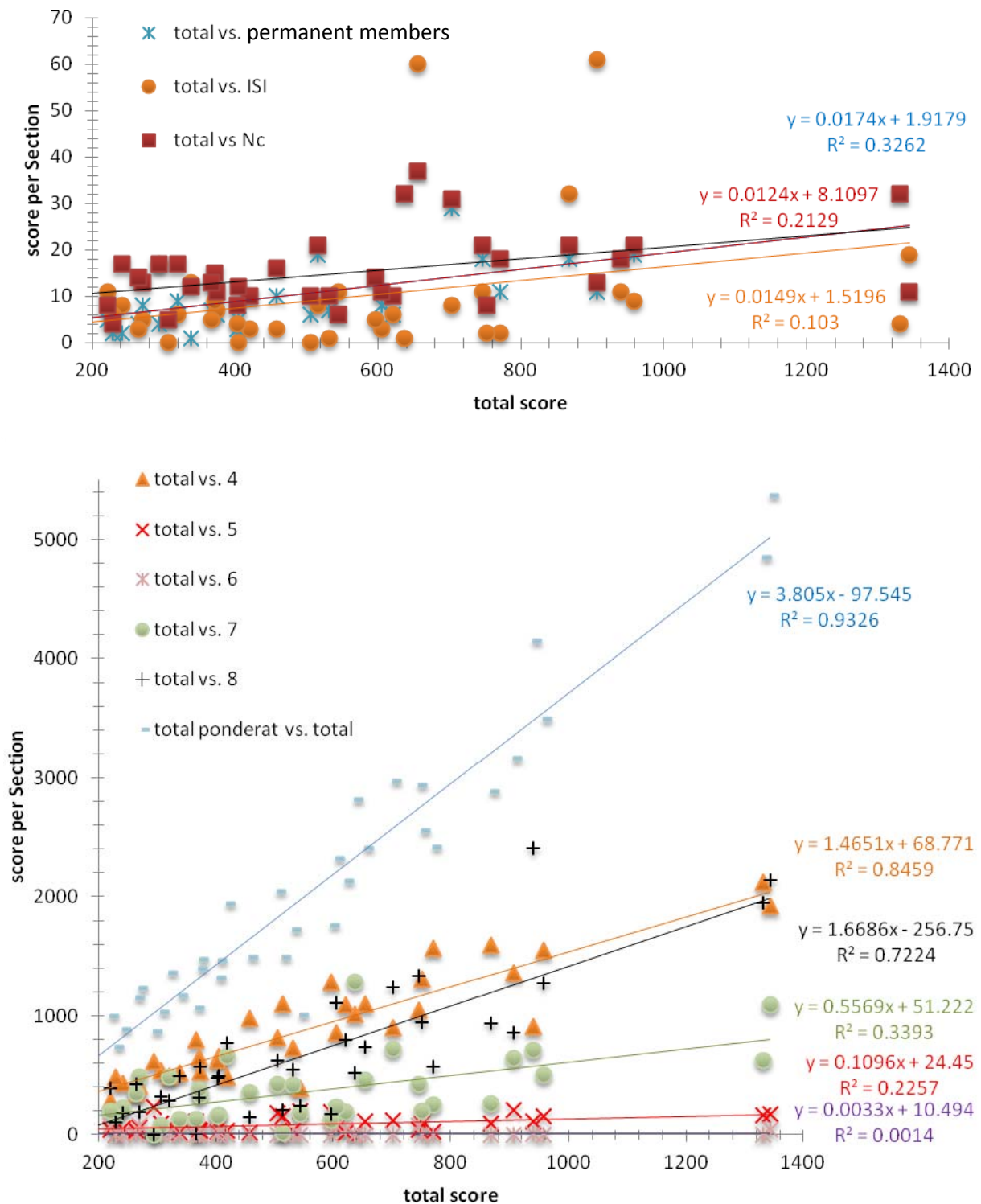


Fig. 4. Statistics for Domains II+III.

We have recently described how ~2% of the employees of the BBU appeared to be responsible for 40% of the research output – and 10% were responsible for 90% of the output, according to criteria such as number of internationally-recognized books (strong presence in WorldCat or similar) and ISI-indexed items (articles, citations, h index).[3] With a very small number of individual exceptions,

these researchers are present in the 56 research units analyzed here, so that the data in the present report may be considered complementary to the data in reference [3]. Of all the data analyzed in either of the studies, one may note that the criterion “number of researchers with $h > 8/h > 4$ within the research unit” (items 4 and 16 in Table 5) is among the easiest to report and to further handle – as opposed to

publication lists, citations, etc. Further still, the sum of the respective h-indices is also a simple parameter to store and handle – as long as a reasonable minimal threshold is maintained (otherwise, a research unit with twenty-five researchers of $h=1$ would arguably excessively surpass one with two $h=10$ researchers). Interestingly, the total score shows a correlation of $R^2=0.76$ with the ‘number of researchers with $h>$ threshold’ as well as with the sums of h-indices (again, counting only those larger than the threshold); the two parameters (number of researchers above threshold and sum of h indices) correlate with $R^2=0.98$. The correlation between total score and sum of h-indices increases to 0.83 when applied only to the research units for which the parameter is different from zero (17 out of 56; 13 of these pertain to Domain I). As in the case of the parameter ‘number of ISI papers’, the correlation is mainly due to the Domain I centers (the overall $R^2=0.76$ becomes 0.82 when applied only to Domain I). As such, the sum of

h-indices above a cutoff appears to be a reasonable predictor of the ranking of a research unit within a classification; in the present study, they range from 9 to 219. Interestingly, although we have previously found individual values of h-indices to be predicted by a Langmuir model implying a maximal value in the range of 33 (i.e., implying that BBU researchers, within the current institutional settings, are particularly unlikely to exceed this limit of h-index), the sum-of-h parameter per research unit seems to fit reasonably to a linear function – and in fact the Hill equation shows weaker correlation (0.61 for all units as opposed to 0.76 for the linear correlation, and 0.82 for only the non-zero-h units as opposed to 0.83). This implies, expectedly, that the research units will find it easier to accumulate increasingly large scores, compared to the individuals – if only by “simple” addition (granted, scientific collaboration between competitive researchers is not “simple” by any default).

Table 6. Key data resulted from the evaluation for the 56 research units; Domain I data are shown on colored background.

Faculty affiliation	F, fraction relative to the maximum score per BBU	Nc	Permanent employees	F per member	F per permanent employee	Section 4	Section 5	Section 6	Section 7	Section 8	ISI papers	Score per member	ISI papers per permanent employee
Political sciences	0.06	17	2	0.1	0.4	432	49	0	222	180	8	14	4.0
-	0.13	6	6	0.6	0.3	386	195	0	177	243	11	91	1.8
Geography	0.24	21	19	0.3	0.1	1551	160	0	505	1275	9	46	0.5
Letters	0.10	10	10	0.3	0.1	484	32	0	655	770	3	42	0.3
Political Sciences	0.16	37	37	0.1	0.1	1096	116	0	462	735	60	18	1.6
Theater and Television	0.17	31	29	0.2	0.1	900	119	0	718	1235	8	23	0.3
Theology, Reformed	0.08	5	5	0.4	0.2	537	94	0	80	320	0	61	0.0
Sports and Physical Education	0.09	15	15	0.2	0.1	653	41	0	385	315	9	25	0.6
Letters	0.15	11	8	0.4	0.2	858	58	70	231	1105	3	55	0.4
Sociology and social assistance	0.08	17	9	0.1	0.1	493	76	20	481	288	6	19	0.7
History and philosophy	0.33	11	11	0.9	0.4	1920	173	45	1088	2140	19	122	1.7
Geography	0.18	21	18	0.3	0.1	1047	47	80	426	1335	11	36	0.6
Letters	0.04	4	2	0.3	0.3	314	16	0	30	210	4	45	2.0
Letters	0.05	14	14	0.1	0.0	254	12	0	138	140	0	13	0.0
Letters	0.19	18	11	0.3	0.2	1564	28	0	252	575	2	43	0.2
Sociology and social	0.07	13	8	0.1	0.1	408	59	80	487	193	5	21	0.6

Faculty affiliation	F, fraction relative to the maximum score per BBU	Nc	Permanent employees	F per member	F per permanent employee	Section 4	Section 5	Section 6	Section 7	Section 8	ISI papers	Score per member	ISI papers per permanent employee
assistance													
Economical sciences	0.23	18	18	0.4	0.2	911	110	10	708	2410	11	52	0.6
-	0.21	21	18	0.3	0.1	1594	97	0	263	931	32	41	1.8
Letters	0.12	10	6	0.4	0.2	813	179	0	425	625		51	0.0
Letters	0.10	12	5	0.2	0.2	660	117	30	168	490	0	34	0.0
History and philosophy	0.10	8	3	0.4	0.4	626	24	0	155	475	4	50	1.3
Letters	0.19	8	8	0.7	0.3	1311	99	0	200	940	2	94	0.3
Letters	0.13	21	19	0.2	0.1	1101	153	10	15	210	8	25	0.4
Political sciences	0.07	17	4	0.1	0.2	617	237	0	0	0	17	17	4.3
Sociology and social assistance	0.06	4	2	0.4	0.3	483	0	0	155	100	6	57	3.0
Political Sciences	0.09	13	7	0.2	0.2	799	114	0	128	10	5	28	0.7
History and philosophy	0.15	14	14	0.3	0.1	1280	189	0	120	170	5	43	0.4
European studies	0.33	32	32	0.3	0.1	2120	161	0	620	1945	4	42	0.1
Theology, orthodox	0.16	32	32	0.1	0.1	1014	0	0	1280	520	1	20	0.0
History and philosophy	0.11	16	10	0.2	0.1	977	13	0	355	145	3	29	0.3
History and philosophy	0.09	11	6	0.2	0.2	530	38	0	331	570	7	34	1.2
Letters	0.13	10	7	0.4	0.2	727	32	0	415	545	1	53	0.1
History and philosophy	0.05	6	6	0.2	0.1	397	4	0	115	130	3	33	0.5
History and philosophy	0.15	10	6	0.4	0.3	1089	14	40	190	800	6	62	1.0
Psychology and educational sciences	0.05	8	5	0.2	0.1	283	45	80	195	393	11	28	2.2
Psychology and educational sciences	0.22	13	11	0.5	0.2	1362	211	0	644	860	61	70	5.5
Psychology and educational sciences	0.08	12	1	0.2	1.0	523	15	0	128	490	13	28	13.0
History and philosophy	0.07	14	4	0.1	0.2	348	34	0	346	420	3	19	0.8
Mathematics and informatics	0.14	13	10	0.3	0.2	1014	38	0	297	535	56	42	5.6
Mathematics and informatics	0.61	18	15	1.0	0.5	4294	168	0	381	3325	242	136	16.1
Chemistry and chemical engineering	0.17	22	16	0.2	0.1	1432	54	54	252	355	79	31	4.9
Biology and geology	0.21	23	15	0.3	0.2	1343	181	25	380	1105	88	36	5.9
Chemistry and chemical	0.20	23	11	0.3	0.2	1496	21	16	493	815	119	36	10.8

Faculty affiliation	F, fraction relative to the maximum score per BBU	Nc	Permanent employees	F per member	F per permanent employee	Section 4	Section 5	Section 6	Section 7	Section 8	ISI papers	Score per member	ISI papers per permanent employee
engineering													
Environment science and engineering	0.15	26	24	0.2	0.1	858	198	89	784	850	71	24	3.0
Chemistry and chemical engineering	0.34	34	7	0.3	0.6	2780	95	0	883	850	138	41	19.7
Chemistry and chemical engineering	0.38	26	8	0.4	0.6	3070	184	20	842	923	174	59	21.8
-	0.07	2	2	1.0	0.4	565	24	20	95	240	33	145	16.5
-	0.21	23	18	0.3	0.1	1795	48	0	238	440	32	36	1.8
Mathematics and informatics	0.06	14	13	0.1	0.1	386	49	0	198	355	10	16	0.8
Biology and geology	0.06	7	2	0.2	0.3	355	40	0	219	300	9	33	4.5
-	0.26	30	30	0.3	0.1	1869	228	315	563	925	136	36	4.5
-	0.76	95	27	0.2	0.3	5874	224	60	1580	2295	281	32	10.4
Physics	1.00	68	50	0.4	0.2	7746	452	21	1370	2682	384	60	7.7
-	0.03	6	5	0.2	0.1	146	14	35	95	290	5	22	1.0
Mathematics and informatics	0.15	23	17	0.2	0.1	1056	70	0	410	565	50	26	2.9
Chemistry and chemical engineering	0.14	6	6	0.7	0.3	1022	39	0	282	595	108	94	18.0

Table 7 illustrates the relative contributions of the 21 major academic structures ("Faculties") of the university to the 56 research units evaluated. Depending on the faculty, the percentage of people involved in organized research activities as evaluated

in the present procedure, ranges from 0 to 100%. No correlations between these percentages and the total scores was found, nor did this percentage depend on the Department, as can be seen in the case of the chemistry research units

Table 7. Affiliations of the evaluated researchers to the Faculties within BBU.

Faculties	Permanent employees members in research units evaluated in 2013	Total faculty members in 2013
Sports and Physical education	15	52
Economical sciences	18	214
Geography	37	81
History and philosophy	60	77
Letters	90	167
Psychology and educational sciences	17	128
Sociology and social assistance	19	43
Political and administrative sciences	50	82
European studies	32	48
Theater and television	29	61
Theology, orthodox	32	27
Theology, reformed	5	14

Faculties	Permanent employees members in research units evaluated in 2013	Total faculty members in 2013
Biology and Geology	17	63
Chemistry and chemical engineering	48	74
Physics	50	41
Mathematics and informatics	55	114
Environment science and engineering	24	33
Law	0	33
Business	0	24
Theology, Greek-Catholic	0	19
Theology, Roman-Catholic	0	11
Not affiliated to a Faculty	106	-

Limitations of the method

We have discussed above how the number of ISI-indexed articles (either directly itself, or indirectly via criterion 3 of Section 4, cf. Table 1) appears to be an excellent predictor of the total score. While this may be taken as administrative advantage (i.e., providing a fast estimate of the total score from just one single simple parameter), it may also be taken to be worrying, as it places exaggerated emphasis on quantity vs. quality. Indeed, it can hardly be argued that an article of impact factor 0.1 should be given equal weight compared to one of impact factor 5. Importantly, however, as detailed below, the total score also correlates with impact factors and h-indices, i.e. with parameters related to quality; this may be explained to suggest that, at least within this institutional setting, groups with larger outputs in term of number of articles also tend to publish higher-impact papers (as judged via impact factors, h, or other parameters). Thus, for larger research units, and over several years, the number of articles does seem a more reasonable descriptor than for individual researchers over shorter timespans. The relative success of various international classification of universities that take into account such parameters as total number of papers is consistent with these observations – although we note that even there the evaluated entity is typically eventually placed in a larger pool of similar entities (e.g., ranked 100-150, or 300-400, as opposed to ranked 123 or 345), notwithstanding the exact numerical calculations that are indeed employed. Nevertheless, it is conceivably possible that the criteria illustrated here would, especially when applied to smaller entities or to individual researchers (for which they were in fact not intended), allow for abuse or distortion via overestimation of the importance of low-impact papers.

An important parameter possibly predictable of the total outcome is the sum of impact factors of the papers. This parameter enters into the total equation in two forms: once directly as such, and once in the form of item 3 in Section 4 (Table 1), where it is added to 10 and divided by the number of authors. The latter formula (which may be criticized in its rather arbitrary choice of using “+10”) is meant to place excessive weight on papers indexed in certain databases, deemed more selective than others (e.g., ISI, ERIH); its limitation is that, within the set of ISI papers, it downplays the importance of impact factor. Yet again, the latter limitation is in itself partly compensated by the inclusion of citation counts, impact factors and h-indices as separate contributors to the score. It is instructive to note that the simple sum of impact factors correlates with the total score with $R^2=0.68$. For Domain I, where impact factors are in general use, the correlation is 0.76. For comparison, the total score is somewhat better correlated with total number of ISI papers: 0.76 for all units, and 0.95 for Domain I. These correlations are then worrying when considering Domain II, where, for many fields of research, impact factors are not defined; on the other hand, it is noteworthy that the second largest total score of the 56 evaluated entities was precisely obtained by such a Domain II research unit, and the trend is further followed throughout the entire list of units, suggesting that the weight of the impact factors can indeed in our protocols be counterbalanced by the humanities-specific criteria.

To give a measure of the caveats needed if attempting to build classifications on the scores yielded by our procedure, one may take the example of two research units with sizeable scores and from the same Domain (the same scientific area, too - chemistry), which for the purpose of this comparison will be denoted CH-L and CH-V. The ratios of their

scores are as follows: 0.9 for total score, permanent members, Section 4, Section 8; less than 0.8 for ISI papers, total impact factor, Sections 5 and 6; 1.0 for Section 7; 1.1 for average impact factor per paper, 1.3 for number of members, 1.5 for sum of h-indices (counting only above threshold), and 2 for number of people with h higher than 8. The higher total score for CH-L is then seen to be due to larger number of papers and larger sum of impact factors (leading to larger score on Section 4) as well as to larger scores for Sections 5 and 8 (funding and prestige). Yet, CH-V presents a 10% higher average impact factor, twice as many large-h members, and 50% larger sum-of-h. Under these conditions, any statement about one unit being generally “better” than the other would be easy to criticize. One may instead argue that, within the institutional setting, where the two units are among the top 5 out of 56, and holding together 10% of the total number of people with larger-than-threshold h-indices, are both at the higher end of the classification and with intrinsic value beyond the classification— and thus arguably “very good” (or, if stronger words may be used, “excellent”). If so, then according to the strict laws of grammar an absolute superlative such as “very” does not allow for any further comparative degrees. If nevertheless one persists in such comparisons, it is important to note that the ratios of the scores will easily oscillate above and below 1, depending on choices made upon criteria of which none is universally recognized as entirely “better” than all others – in which case any such forced hierarchy between two closely-matched entities will be the result of administrative or personal choices, rather than a manifestation of hard facts/reality (even though, undeniably, both units discussed in this example, CH-V and CH-L, are far exceeding the majority of the others on several criteria as well as on the total score).

The degree to which the total scores for the 56 units, especially in Domain 1, correlate with number of papers as well as with h-indices (which in turn are controlled by citations, and indirectly by impact factors), may be taken to imply two things. First, that these parameters are good predictors of the total score – and that hence a top research unit can already be identified and set aside based on a preliminary screening of, e.g., only the h-indices of its members. Yet, we have previously warned[3] that the h, like the number of papers and the impact factors, are heavily domain/field-biased. As such, the set of criteria described here may be criticized for favoring chemists over e.g. geologists or mathematicians. If, however, one would set further-differentiated thresholds for each discipline, and use these for setting apart “very good” individuals – or indeed

“very good” research units led by these individuals, such preliminary screening could well save significant amounts of time on the part of evaluator as well as evaluated units. Yet again, abuse of the h-index may lead to intriguing ethical issues such as: is the h=15 of a postdoc publishing for several years with a high-impact author equally valuable as the h=15 index of a truly independent principal investigator, or equally valuable as the h=15 of a postdoc who shares his papers not with one but with three other people?

On a further technical note, one may argue that the funding taken into account for Section 5 (data in Table 2) should not be used as such, but rather emended to only take into account certain types of expenses – e.g. salaries of temporary researchers (or indeed number of research positions newly created by the grant), consumables, or others that would be reasoned to hold a more directly-accountable value for the respective research unit. In favor of using the simpler approach of Table 2 one expects the relative percentages of various expenses to be largely similar and controlled by regulations; also, it may be argued that the manner in which the funds are distributed within the research unit is, while itself important, a subtlety that deserves a full separate discussion.

Yet another point for potential critique is the inclusion of patent applications among the criteria: one may argue that any “application” has by definition a certain chance of being rejected, and thus rendered valueless, much in the same way that manuscripts may be rejected by journals. In favor of the criterion remains the fact that patent applications are processed in quite different manners than manuscripts and, at least those from BBU, enjoy a reassuringly high rate of success. Also, it may be argued that the relatively low number of patents across the BBU justifies a range of efforts to stimulate an increase.

The points awarded for ISI-indexed proceedings by making recourse to the impact factor have also caused some critique, since the term “Proceedings” was too loosely defined. A Proceedings by definition is not a journal and has no impact factor; yet some journals will accept to publish papers in special volumes dedicated to conferences, thus associating the prestige of the journal with the conference. It would be important for the future to more clearly delineate the situations – and also delineate between abstracts and full papers published in such “proceedings” special issues.

The points awarded for books and book chapters may be argued to be excessively low by comparison with simple articles. For instance, two articles in a journal of impact factor 0.1 would receive more

points than an entire book at any of the most reputed international publishers. A substantial increase of these book-related criteria (e.g., ten-fold, for those with high impact) may thus be proposed – alongside perhaps with a closer scrutiny of the nature of the book, so as to more efficiently screen off the lower-impact ones (e.g., by verifying their presence in a substantial number of libraries).

Overall, it is important to note that the present report seeks to describe a state of facts, as to which criteria were already used, and to what administrative facts they already led, rather than to propose a complete solution. We have commented upon how this state has its limitations, on points where it can be criticized, and on points where it can be improved. The BBU is committed, via its official documents, to such improvement by the next evaluation cycle, and it is hoped that the present account may be a useful reference in that respect.

Conclusions

The evaluation procedure allowed the University to update its database of research output, and to help reinforce, both with the researchers and with the management, the types and quality of output desirable in the future. The need to further develop applied research, the need to reorganize some of the lesser-performing research units, and the need to selectively support the higher-ranked units (whether this rank be in terms of overall score or in terms of score per member), were among the useful conclusions. The numerical data has not been, up to this point, further employed in taking any administrative actions; all of the 56 research units received a “Certificate of accreditation” at the end of the procedure. The Scientific Council and the Senate agreed to further discuss the criteria, and apply improved versions to any future evaluation procedure (the next schedule procedure being proposed to occur in 4 years’ time). A key issue was that the academic body, and especially the Senate of the University, saw their confidence in the evaluation boosted by the fact that any readily-quantifiable criterion is essentially at least nominally incorporated in the system. If this advantage is to be given up, one could probably focus the evaluation on simply the number of research articles, and add to this Section 6 (innovation, application) – probably weighted by an extra order of magnitude so that it can make a straightforwardly interpretable difference.

Acknowledgements

The Romanian Academy is thanked for sharing algorithms that were applied in the evaluation scheme described here. The members of the Scientific Council of the Babeş-Bolyai University, the two bodies of evaluators (and especially their coordinators, Profs Maria Roth and Simion Aştălean), and Prof Daniel David are thanked for helpful discussions. This paper is dedicated to Nicolae Marina.

Supporting Information available: electronic forms for Domains I, II and III.

References

- [1] Ad-Astra, http://www.ad-astra.ro/cartealbalba/domains_authors.php (2012).
- [2] I. D. Caruntu, [Romanian Universities performances and Shanghai classification], *Rev Med Chir Soc Med Nat Iasi* 110 (2006) 257-8.
- [3] R. Silaghi-Dumitrescu, A. Sabau, Scientometric analysis of relative performance in a key University in Romania, *Scientometrics* 99 (2014) 463-474.