hal-00664539, version 1 - 30 Jan 2012

Are We Overdoing It with the Hirsch-Index H ?

Elemér E Rosinger

Department of Mathematics
and Applied Mathematics
University of Pretoria
Pretoria
0002 South Africa
eerosinger@hotmail.com

Dedicated to Marie-Louise Nykamp

Abstract

It is shown that measures of complex entities or performance given by one single number, such as the IQ, or the lately fashionable Hirsch-Index H, are highly questionable not least due to the Principle of Increasing Irrelevance of Preference Type Information.

> "... creativity often consists of finding hidden assumptions. And removing those assumptions can open up a new set of possibilities ..."
>
> Henry R Sturman

> "History is written with the feet ..."
>
> Chinese Ex-Chairman Mao,
> of the Long March fame ...

hal-00664539, version 1 - 30 Jan 2012

In the latest arxiv:1201.5476 paper, [3], it is argued based on a study of papers in astrophysics that the mean relation $H = 0.5(1 + \sqrt{c})$ holds, where c is the total number of citations of a researcher. This implies that the recently fashionable Hirsch-Index H does not in fact seem to be a significantly measure.

So much for the ongoing efforts to come up with better and better "bean counting" methods for the evaluation of the relevance and importance of research output in science.

What should in general be an elementary commonsense fact about the inherent limitations of "bean counting" seems time and again to be forgotten by the so called "management" of science. This long ongoing amnesia comes with the lately growing popularity of the Hirsch-Index, or in short H-index, which its adepts believe to be an appropriate and rather simple numerical measure of a researcher's visibility - and thus, relevance - in science.

There should, however, be not much surprise in seeing such a numerical approach emerging now and then, and gaining a wider popularity for a while. One of them for instance, the measuring of the IQ-index of individuals, has for about a century by now gained a considerable credibility in the widest circles. And no argument against it seems to be able to point the eyes of its enthusiastic supporters to what should otherwise be easily seen as its ridiculously simplistic and inadequate nature.
Among such utterly disregarded arguments happens to be the following one as well. In earlier times, at female beauty contests only the physical appearance of the contestants was considered, and they were not expected to answer any question coming from the jury. Yet at such superficial beauty contests no less than six numbers associated with the respective females were considered as particularly important, namely

> age, height, weight, chest circumference, waist circumference, and bottom circumference

On the other hand, IQ is supposed to be given *not* by six, but by

hal-00664539, version 1 - 30 Jan 2012

one single number, although it is supposed to measure satisfactorily human intelligence which, as far as we happen to know, it is by far the most complex and subtle entity on Planet Earth, an entity certainly immensely more involved and multifaceted than mere superficial female beauty.

And then the question arises :

> What may be the reason that we keep adopting time and again such ridiculously primitive measures given by *one single number* like IQ when it comes to human intelligence, or for that matter, like the Hirsch Index H when measuring the quality and importance of the research output of scientists ?

Well, quite likely, the psychological dynamics which leads to such an utter failure, and on top of it, to the ongoing refusal of recognizing that failure, may be as follows :

> 1) We clearly understand from the very beginning that the entity we intend to evaluate is indeed complex, subtle, involved, multifaceted, and so on.
>
> 2) Equally clearly, however, and no less importantly, we do not in any way whatsoever intend to get involved in any kind of measuring process whose inevitably difficult conceptual structure would be near enough to the involved nature of the entity which is the subject of evaluation, a nature as mentioned in 1) above.
>
> 3) And then, based on a hubristic self-confidence in our cleverness, we decide to produce no more than one single number as the measure of that entity.
>
> 4) Indeed, our top priority is not so much to make justice to the entity subject to measurement, as rather to turn that measuring venture into a comfortable event.

hal-00664539, version 1 - 30 Jan 2012

> 5) An when it comes to enjoying one's comfort, and bask at the same time in the assumed success of one's cleverness, well, no any kind of arguments are supposed to be allowed to dislodge one from such a state.

Running, therefore, the risk to discomfort a larger number of supporters of ever more clever ways of "bean counting", let us briefly point to the following which, even if refused to be considered, are nevertheless relevant at least as a *warning*.

Both human intelligence, as well as a researcher's visibility - and thus, relevance - in science, are entities which involve a considerable amount of important features. And nowadays, this amount is in fact not yet anywhere near to be clearly known by anybody. Furthermore, it is even less knowledge about the specific aspects of each and every feature of the mentioned kind of entities. The inevitable and critically important effect is that, whenever trying to compare the intelligence of two persons, or for that matter, the relevance of the research output of two scientists, we are faced with what may be called a "multiple dimensional preference choice", or in short, MDPC, a choice in which we do not know even the number of dimensions involved, let alone the nature of the features upon which we are supposed to make choices.

And to give an idea about what may indeed be involved here, let us recall the discipline within Optimization which goes by the name of MCDM, that is, "multiple criteria decision making". Briefly, the basic problem dealt with by MCDM is as follows. A given decision maker has a set X of choices and a number $n \geq 2$ of well defined criteria $c_1, \ldots, c_n : X \longrightarrow \mathbb{R}$ which are usually *conflicting* among them to a lesser or greater extent. His or her problem is to find $x*$ in X, such that it *maximizes simultaneously* all the criteria $c_1, \ldots, c_n$.
Of course, the paradigmatic case of that situation is described by the well known adage :

> One can't have one's cake, and eat it, too

where the number of conflicting criteria is merely $n = 2$, thus the smallest possible to allow a nontrivial situation, yet the problem is

hal-00664539, version 1 - 30 Jan 2012

famously difficult, as common wisdom has known it for ages.

Now, one of the novelties regarding MCDM is that, the larger the number n of criteria, the smaller the relevance of pairwise comparisons of decisions x and y in X, when trying to obtain the best decision $x*$ in X, [2].

This is formulated as the "Principle of Increasing Irrelevance of Preference Type Information", or in short, PIIPTI, [2].

And as it turns out, the relevance of preference type information in the MCDM context does in fact decreases *exponentially*, as the number n of criteria increases. This phenomenon is but a direct consequence of a simple property of the surface of a sphere in higher dimensional Euclidean spaces, a property which may intuitively be described as follows : it is not worth buying higher dimensional water melons, since more and more of their content will be in their shell, [1].
But to put it in more scientific - this time, physics - terms, the very phenomenon of temperature is an immediate consequence of that higher dimensional geometric property, [1].

The obvious relevance of PIIPTI in pointing to the utter inadequacy of IQ or of the Hirsch-Index H is as follows.

> 1) Clearly, neither IQ, nor H are the results of any optimization. On the other hand, both of them are claimed to measure by one single number a simply unknown number of important features of one human's intelligence, respectively, of the quality and relevance of the research of one scientist.
>
> 2) In this way, from the start it is obvious that one is not supposed to know even the number $n \geq 2$ of entities which may be relevant in measuring what IQ and H are supposed to measure.
>
> 3) Further, even if that number n of the respective entities - a number which no doubt, is considerably larger than 2

hal-00664539, version 1 - 30 Jan 2012

- would be clearly known both for human intelligence, respectively, the quality and relevance of a research output, one would inevitably be involved in analyzing the important features of each such entities, in order to be able to measure them in any conceivable way.

4) Last, and not least, such measurements must involve comparisons between the mentioned entities, and within each such entity, between its relevant features. Thus inevitably, the action of PIIPTI appears, since the number of entities, and within each of them, the number of relevant features - both kind of such numbers still completely unknown at present - can be significantly larger then 2.

So much for the adequacy of "bean counting" as an important tool in "science management" ...
And if you can't beat them, well, then you better join them ...

Long live, therefore, the Hirsch-Index H ...

References

[1] Manin, Yu I : Mathematics and Physics. Birkhauser, Boston, 1981

[2] Rosinger E E : PIIPTI, or the Principle of Increasing Irrelevance of Preference Type Information. arXiv:math/0506619

[3] Spruit H C : The relative significance of the H-index. arXiv:1201.5476