

The background features a large, stylized eye shape in the center, composed of a dark iris and a white pupil. This eye is overlaid on a complex grid of white lines that form a perspective of a tunnel or a series of receding planes. The overall color palette is dominated by shades of purple, blue, and white, with a soft, ethereal glow.

“Communicating Science and Publication Cultures In the Humanities”

LabCom, UBI
PTDC/CCI-CIN/119869/2010

The Evaluation of Research by Scientometric Indicators

Péter Vinkler

Péter Vinkler

(1941 -)



- Graduate Chemistry, Univ. of Szeged, Hungria, 1966
- PhD *Spectroscopy of Infrared*, 1974
- Made Doctor of the *Hungarian Academy of Sciences*, 2003
- Scientific advisor / Scientific secretary do *Chemical Research Center (CRC)*, Budapest
- Untill Novembro 2009: Director of the *Scientific Publication Data bank* of the *Hungarian Academy of Sciences (HAS)*.
- Member of the *American Society of Information Science and Technology* and of the *International Society of Scientometrics and Informetrics*
- Editorial Advisor to *Scientometrics*. In 2009 he was awarded the *Derek de Solla Price Medal* for his contribution to the quantitative studies of sciences

I – Assumptions / definitions

- **Science**: produce and communicate scientific knowledge. (...) «Science is a socially shared and socially validated body of knowledge.” (p.1)
- **The essence** of scientific research is **communication**.
(F. Crick,1979)
- **Scientometrics**: analyzes quantitative aspects (generation, propagation, utilization...) in order to contribute to a better understanding of the mechanism of scientific research activities (Braun *et al.* 1987)

I – Assumptions / definitions

- “**Yes, but...**”: Scientometrics not belongs primarily to a particular scientific discipline. Its goal is one more efficient management of science (*science's policies*)
- May be «**the science of science**», but not as philosophy or epistemology (involves theoretical and empirical studies)
- **Evaluative bibliometrics** (Narin, F., 1976), as measuring quantitative indicators, based on previous publications, is just a field within **scientometrics**, wich **represents a broader view** (non bibliometrical data, human capacity, grants, equipments, etc..)

I – Assumptions / definitions

– How to develop an evaluative scientometrics?

a) Quantitative measures of research are only partial indicators

b) The most fruitful approach is the combined use of multiple indicators

– V.g, bibliometric assessment by ‘peer review’ must be assumed as ‘goods’, i.e. having scientific *value* and *use value*: originality, validity, brightness, generality, coherence, applicability, ‘referentiality’...

I – Assumptions / definitions

- Mas se não se levar em conta, as diferentes situações à partida de cada investigador, equipa, etc., a avaliação pode ser injusta. “The are no absolute scientometrics methods for assessing scientific eminence.” (questão do mérito absoluto)
- Assim, deve-se encontrar os melhores indicadores e métodos para determinar a *posição relativa* de investigadores, equipas, instituições, países, regiões, ...
- Importância cada vez maior dos composite indexes (para avaliar convergentemente diferentes aspectos)

I – Assumptions / definitions

- Maior paradoxo da *scientometrics* é que, estatisticamente, os métodos só são válidos “for a large (whole) sets of publications, but they are applied for assessing part-sets”.
- Importância cada vez maior da “microscientometrics” (ao nível de equipa, laboratório, especialidades, ...). «Each assessment may be regarded as a special exercise. There are no recommendations detailed (...); there are general relationships and conclusions...» (p.4)

I – Assumptions / definitions

«Science policy-makers [both institutional / national level] demand more than scientometricians can offer.» (p.4)

Por isso, os *scientometricians* têm de tomar muita atenção à aplicabilidade, escalabilidade, limitações, incertezas, etc., ainda que o princípio *aestimare* *necesse est* seja imprescindível nos diferentes níveis.

I – Assumptions / definitions

– A *scientometrics* não é nem uma “*hard science*” nem uma “*soft science*”, mas um misto cuja relação é feita pela Estatística, a qual vale em primeiro lugar para conjuntos. Os seus resultados “**may be regarded as trends rather than strict rules**”.

– Interdependência dos indicadores: *GF*, *RPG*, *CCI*, *PS*, *h-index*, *π -index*, *TotalPublicationProductivity*, etc., não devem ser usados separadamente.

I – Assumptions / definitions

– «The basic assumptions of evaluative scientometrics – the information unit of sciences is the scientific paper, and the unit of impact is citation – are only crude, statistical approximations.»

n.b.: Mas é com estas unidades básicas que o autor trabalha sempre ao longo do livro!
(p. 9, *et alibus*)

I – Assumptions / definitions

«It is clear that any publication assessment method **must cover** the *amount* of scientific information produced by the scientists evaluated, the *eminence* of the publication channels used, and international *acknowledgment* of the results published.» (p. 5)

I – Assumptions / definitions

— «The numbers of publications and citations of individuals, laboratories or countries are easily available via international data banks [e.g., SCI ISI, Thomson Reuters]. These numbers resemble a gun on the stage of a theatre, which must be sooner or later be fired.»

N.B. — Necessidade de formação e experiência em *Scientometrics*. Necessidade de equipas especializadas nas Universidades, para poderem fazer a aplicação *in loco*.

SCI - Science Citation Information; ISI - Information Sciences Institute; ISI - Institute for Scientific Information



II - Basic Categories of Scientometrics


II – Basic Categories of Scientometrics

Manual for “Good Scientometrics Practice”
(Chicago, 1995)

- *Scientometric aspects* (características quantificáveis de coisas e fenômenos na ciência)
- *Scientometric elements* (autores, artigos, revistas, referências, citações...)
- A *scientometric set* (conjunto de artigos de uma revista, GF das revistas num campo, etc.)
- A *scientometric measure* (valor numérico simples, ou algoritmo que caracteriza aspectos: número de artigos, revistas, citações, rácio de citações por artigo, outros rácios, ...)
- A *scientometric unit* (artigos, citações; artigos/investigador, artigos/citações, ...)

II – Basic Categories of Scientometrics

- *Scientometric system* (entidade à qual vários ou muitos elementos cientométricos são atribuídos: equipas, revistas, campos de investigação, países, regiões...)
- *The scientometric indicator* (medida que pode caracterizar um ou vários aspectos da investigação atribuída a um sistema cientométrico. Exs.: GF de uma Revista num ano; produtividade científica de uma Universidade, etc.)
- *Evaluative scientometrics*: campo especial da *scientometrics* aplicada em sistemas que visa AVALIAR, i.e., tirar conclusões quantitativas sobre a *performance* de uma organização; estudos comparativos sobre produção, disseminação da informação, aferidas por um *scientometric impact* (v.g. um autor aplica um novo método antes publicado; uma patente farmacêutica refere um artigo sobre a síntese de um novo componente, ...).



**III - Classification of the indicators of
evaluative scientometrics**

III - Classification of the indicators of evaluative scientometrics

As unidades temáticas e organizacionais (campos, sub-campos, disciplinas, ... ; regiões, países, universidades, institutos, laboratórios, departamentos, equipas e investigadores), os objectos, os fenómenos, os bancos de dados, as fórmulas, etc., devem ser classificados segundo indicadores que notam aspectos comuns e aspectos não-comuns.

Tais indicadores podem classificar

- Actividade
- Produtividade
- Progresso

e devem distinguir-se segundo

- Qualidade (exactidão, elegância, generalidade, ...)
- Importância (potencial impacto na ciência, sociedade...)
- Impacto (influência actual)

III – Classification of the indicators of evaluative scientometrics

Os indicadores avaliativos mostram os

- *inputs*
- *outputs*

e são

- **Funcionais** (referem-se a objectivos da investigação científica)
- **Instrumentais** (aspectos metodológicos)

Os Índices podem ser:

- *Quantitative* ou
- *Qualitative*

- *Databased* ou
- *Perceptual* (*constructed by peers or committees*)

III – Classification of the indicators of evaluative scientometrics

Os indicadores podem ainda ser classificados de acordo com os elementos de um conjunto a que se referem:

- publicação
- citação ou referência
- patente
- potencial (capacidade humana, bolsas, equipamentos ...)

E conforme o nível de avaliação, são:

- micro
- meso
- macro indicadores

Em relação ao tempo, são:

- Dependentes ou independentes do tempo;
- Relacionam-se ou negligenciam a relação com o tempo.

III - Classification of the indicators of evaluative scientometrics

— *Gross indicators*: medem um único conjunto cientométrico, ou apenas um único nível hierárquico: quantidade bibliométrica (n° publs. de uma org. em Δ/t), capacidade de investigação (n° investigs., equipas, faculds., Δt , ano), € investido em instrumentos em Δt ; Orçamento anual; total de citantes de grupo de publicações, n° de citações, n° de artigos, etc. [são os mais usados para comparar uma mesma instituição em Δt e $\Delta t'$ (aumento, diminuição; ou instituições comparáveis)]

Os que medem mais do que um conjunto ou nível são chamados *complex indicators*: aplicam “reference standards” com base em equações complexas, e podem ser qualificados de “*Especific*” (GF; n° cit/n°papers; n°cit./n° citers; n° cits/ Δt /investigador(es), etc.: para comparar orgs., áreas, etc.), “*Balance*” (rácios complexos: v.g., n° de cits. de uma rev. noutra, em relação com as cits. que a 1ª faz da 2ª; ou n° de revs. que citam outra, em relação com as cits. que esta faz daquelas; comparações mais finas), “*Distribution*” (v.g., n° de artigos cit. 0, 1, 2, 3, 4, 5, etc.; 15+citds num ano; n° revs. citads / n° artigos, etc.), “*Relative*” (OCR - Observed Citation Rate / ECR - Expected Citation Rate; média de cits. do artigo de um autor face aos *papers* cits. no SCS - Standardization Citation Score; Somatórios dos GFs das revistas onde os artigos (P) foram publicados, relacionados com o GF de subcampos correspondentes x pelo n° de P) e “*Composite indicators*”: relações altamente abstractas, indicadores de indicadores, ou de sub-indicadores (p.21).

Por isso...

“Eu ‘pecador’ me confesso...”

(equações para normalização de índices)

6. Indicators above or below the mean

$$I_{qc}^t = \begin{cases} 1 & \text{if } w > (1 + p) \\ 0 & \text{if } (1 - p) \leq w \leq (1 + p) \\ -1 & \text{if } w < (1 - p) \end{cases}$$

where $w = x_{qc}^t / x_{qc=\bar{c}}^t$

7. Cyclical indicators (OECD)

$$I_{qc}^t = \frac{x_{qc}^t - E_t(x_{qc}^t)}{E_t(|x_{qc}^t - E_t(x_{qc}^t)|)}$$

8. Balance of opinions (EC)

$$I_{qc}^t = \frac{100}{N_e} \sum_e \text{sgn}_e(x_{qc}^t - x_{qc}^{t-1})$$

9. Percentage of annual differences over consecutive years

$$I_{qc}^t = \frac{x_{qc}^t - x_{qc}^{t-1}}{x_{qc}^t}$$

Note: x_{qc}^t is the value of indicator q for country c at time t . \bar{c} is the reference country. The operator sgn gives the sign of the argument (i.e. +1 if the argument is positive, -1 if the argument is negative). N_e is the total number of experts surveyed. P^i is the i -th percentile of the distribution of the indicator x_{qc}^t and p an arbitrary threshold around the mean.

III - Classificação dos Indicadores de Avaliação Cientométrica

Os complex indicators podem ser rastreados/reconduzidos aos gross indicators:

V.g., para calcular a Relative Subfied Citedness (RW) temos de ter em conta as citações [conjunto A] obtidas pelos artigos efectivamente avaliados [conjunto B].

A “mean citedness” (GF_m) de um conjunto de revistas (*‘meta-journal’*, *‘meta set’*) tidas como *standard* pode ser calculada a partir das respectivas citações [conjunto C] e artigos [conjunto D].

Estes indicadores usam-se para caracterizar a actividade global das organizações e sistemas (Univs., Países, Mundo), e devem ser escolhidos os mais apropriados para os diferentes aspectos dos diferentes sistemas cientométricos.



IV - Publication growth in science

IV – Publication growth in science

- 1º Modelo cientométrico de crescimento científico: **a)** período inicial de lento crescimento, **b)** fase de crescimento rápido e exponencial, **c)** desacelaramento do crescimento, aproximando-se de um limite superior (Price 1963)
- 2º Modelo: “Sistema organizado de conhecimento” (Kant). “Scientific research” alimenta-se de informação *inputs* que geram *outputs* (new or restructured knowledge)
- 3º Modelo (mais usado *in scientometrics*): crescimento cumulativo de publicações. O número cumulativo de publicações num dado ano depende das publicações no princípio do ano, o rácio de crescimento, e o lapso de tempo decorrido. É um modelo linear, rapidamente exponencial.

Segundo um logaritmo desenvolvido por Schummer (1997) para a Química, o modelo linear dá um rácio de crescimento de 5.5%/ano e duplica a cada 12.9 anos.

Mas este modelo é inapropriado se não leva em conta “the ageing of information”. Além de que “science works with great redundancy.” (p.25) Estudos históricos (na Física, Geologia e na Química) mostram que o modelo linear só encontra verificação em certos períodos de tempo, e não tem qualquer valor preditivo.

IV – Publication growth in science

RPG - ‘Relative Publication Growth index’ foi introduzido para caracterizar o crescimento e envelhecimento relativo da literatura científica (+ Chem., Phy., Biology e Maths) e determinar o MAPR: Mean Annual Percentage Rate.

– Unidade básica: “**número de artigos de revista publicados numa dada unidade de tempo**” (ano, de preferência) – RPG(t) relaciona o número de publicações de um dado ano com o número publicado num período precedente seleccionado (2, 5, 10 ou 20 anos, de preferência). A extensão de t deve assumir o RR (*maximum age of recent & relevant papers*). [RRK- *Recent & Relevant knowlegde: papers 0-20 anos*]

– RR obtém-se pelo número total de referências; mas devem expurgar-se duplas referências (v.g., um artigo citado 20x noutra, só vale 1 referência) [equação pp. 28ss].

– No entender do autor, é possível usar funções altamente sofisticadas para avaliar o crescimento da informação científica, mas o MAPR / RPG exprimem com precisão a dinâmica do desenvolvimento científico.

– Os resultados de P. Vinkler referem-se praticamente só à [Química](#)(!) e sub-áreas, onde, segundo o autor, são precisos. [Extrapoláveis para as SC / A&H? Não.]



V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

V – Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

Relevance, Validity & Applicability to characterize the scientific eminence of journals

- **Relevance**: overlap of the information with the interest of the reader
- **Validity**: reliability of the results (methods, data relationships, conclusions)
- **Applicability**: the potential use of the information published

«Scientometrics **cannot measure the inherent value or impact** of information directly. Nevertheless, it can be used indirectly, **via citations**.» (p.35)

The Citedness!

V – Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

The idea of *Citedness*:

- Gross & Gross, 1927 (para a Química: JAQS)
- Raisig, 1960 – “citation is Research Potential Realized” (**RPR**)

$$\mathbf{RPR} = 0 \quad \underline{\vee} \quad \mathbf{RPR} = 1$$

(mas para Raisig a citação de um artigo de uma revista noutra só conta uma vez, mesmo que nela seja citado mais vezes)

- F. Narin, 1976 (“evaluative scientometrics”)
- E. Garfield, 1976, 1979, 1983, 1994, 1998... (IF → GF, Vinkler, 2000)

Citedness

GF = n°cits. ano y / \sum n°total de artigos de $y-1$ e $y-2$, ie., 2 últimos anos. Ex.: 901 [cits. 2010] / 301 [153 arts. 2008 + 148 arts. 2009] → **GF**=2.993

V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

Standardised description of the Garfield Factor

As suggested above, the citedness of journals calculated by Equation (5.2) should be termed as the *Garfield Factor*:

$$GF_y = \frac{C_y}{P_{y-1} + P_{y-2}} \quad (5.2)$$

where GF_y is the Garfield Factor of a journal in year y , C_y is the total number of citations (including all types of self-citations) received in year y by the papers (articles, reviews, letters, notes, etc., i.e. 'citable items', see Chapter 11) published in the journal in years $y - 2$ and $y - 1$, and P_{y-1} and P_{y-2} are the number of papers published in the respective years.

V – Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

Chances for *citation/citedness* [CC]

– Menard's law (1971)

- n ^o papers:	1,	10,	20,	50
- % cit.:	6%,	50%,	90%,	100%

– Gomperz (1969), Price (1970), Glänzel *et al.* (1983), Rosseau (1992), Plomp (1994), Baldi & Hargens (1997), Redner (1998), etc..

– Lehman *et al.* (2003) - citadness by previous classification:

- Phenomenological: > Prob. de CC
- Theoretical
- Experimental
- Instrumental (*or reviewing*) < Prob. de CC

– Vinkler (2004): criou um fórmula de CC que conclui: «the GF of journals may be assumed to provide a measure of *chances for citedness*.» ('**but only** by field', não de uma revista concreta; + em *Natural Sciences*; > Chemistry! cf. p.49-50; p. 268: "physics and chemistry are overrepresented"!) «The increase in CC ou GF is naturally valid only for *the mean value of a whole set of journals (i.e. meta-journal) of a field or subfield.*»

N.B.: isto não se aplica em SC e A&H; não há dados suficientes! (Glänzel, Porto, 8th May 2012)

V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

The **GF** and the Relative Publication Growth (RPG)

Table 5.3

Number of papers (*P*) and references (*R*) and Garfield Factors (*GF*) of the *Journal of the American Chemical Society* and *Journal of Organic Chemistry* between 1981 and 2005

Year (y)	<i>J. Am. Chem. Soc.</i>					<i>J. Org. Chem.</i>				
	GF	P	R	C	R/P	GF	P	R	C	R/P
1981	4.259	1879	57,127	15,206	30.40	1.942	1271	32,625	4685	25.67
1985	4.319	1834	62,512	15,882	34.09	2.155	1299	37,482	5189	28.85
1990	4.463	1992	72,938	17,236	36.62	2.599	1162	41,954	6308	36.10
1995	5.263	2137	87,039	23,210	40.73	3.251	1305	51,287	8501	39.30
1998	5.725	2279	92,671	25,144	40.66	3.500	1562	62,046	10,588	39.72
2005	7.419	3391	134,527	45,707	39.67	3.675	1603	73,019	10,834	45.55
Increase in per cent										
1998/1981	34.42	21.29	62.22	65.36	33.75	80.23	22.90	90.18	126.00	54.73
2005/1981	74.20	80.47	135.49	200.59	30.49	90.26	26.12	123.81	131.25	77.44

C: number of citations obtained in year *y* to papers published in *y* - 1 and *y* - 2 years.
Source of data: Journal Citation Reports (JCR), Science Edition, Thomson Reuters.

Persons *et al.* (2004): «the relative increase of references, authors and publications depends linearly of the numbers of co-authors. The increase could be attributed to the increase in the number of authors per paper»! (p. 44)

V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

– «The majority of researchers assumes that the **GF** represent the mean citedness of papers in a given journal»

– n.b.: *the bias is well-known* (Seglen, 1992) por isso alguns relativizam as CC (Amin & Maabe, 2000); Vinkler, p. 50, relativiza o enviesamento: reorganização das revistas, alteração editorial, etc.)

[n.b.: journals with low or extremely low n^o papers= high GF]

– Por isso, Vinkler não hesita: «the citation may be considered as the scientometric unit of the international impact of scientific publications.»

– Mas nota que os 'relative scientometric indicators' (GF, CC, RPG, CCI - Current Contribution Index, SIC-Specific Impact Contribution, ...) valem para um 'set of journals' [meta-journal] nunca para uma só revista.

– **SCI - Science Citation Index, JCR - Journal Citations Report**

V – Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

Relation between quantity and quality

- Correlacionar *citedness* (C/P) com n^o de *papers* (P) e retirar daí 'quality' (?) [N.B.: here 'quality' = 'impact'!]
- Há estudos para todos os gostos: uns que não encontram correlação forte: Kim&Kim, 2000, ...; outros sim. Seguindo Pearson, Vinkler afirma: «**a higher number of papers may attract a higher number of citations**». (p.52) [i.e, para o autor, o GF é sinónimo de qualidade]
- Cole&Cole (1973), classifica os autores em: *perfectionist, prolific, silent and mass producer*. Com estas categorias, Feist (1997), correlacionou «quality and eminence, depending on quantity».
 - perfectionist :C/P=1.63
 - prolific: C/P=0.72
 - silent: C/P: 0.25
 - mass producer:C/P=0.12

[Quadro p. 53. A correlação não é directa, mas tematização epistemologicamente muito fraca!!!]

V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

CCI?

- CCI: «the measure of contribution to the total impact of journals in a field, made by a given journal’ is in direct correlation with the number of citations, if the impact is measured by the total number of citations received.»
- «The **CCI** may be assumed to provide an indicator of journals representing the current impact of information published at any time by the journal, on generating new information in the correspondent field.»

V - Scientific eminence of journals

The **G**arfield **F**actor and the **C**urrent **C**ontribution **I**ndex

The π_v -index

– É um índice recente que visa avaliar apenas a importância de um **“elite set papers”** (*highly cited papers / most influential papers*) para determinar **“the eminence”** [i.e, “scientific standing’] de certas revistas.

Os *papers* são classificados por ordem decrescente de n^ocits., e mediante uma equação [(10 log P) - 1, onde P é o n^o total de artigos], obtém-se o “elite set”. É um índice muito constante quando aplicado a períodos longos (10-15 anos) e muito sensível a qualquer mudança de quantidade de cites. no ‘set’.

Concluding remark

«There are, at present, several indicators and methods in the scientometric literature for characterizing the eminence of scientific journals. **None of these, however, may be regarded as correct from all aspects.**»



VI - The ageing of scientific information

VI - The ageing of scientific information

— «Most scientific information becomes obsolete with the elapsed time. This process can be described as a function of years and citations.» (p.63)

— Vários autores (Burton & Kleber, 1960; Braun *et al.*, 1987; Gupta, 1990, ...) *equacionaram* esta função de modo a tentar obter uma espécie de «tempo de vida médio» da informação científica.

— Questão: é o «tempo de vida médio» que tem vindo a diminuir ou isso é antes o resultado do crescimento exponencial de revistas / *papers* / autores? (Egghe and Rousseau, 1990).

VI - The ageing of scientific information

Methods:

A ideia principal é medir o número de citações recebidas por um *paper* em anos consecutivos depois do ano da publicação, e/ou comparando um determinado ano /período de tempo com outro ano/período de tempo, de modo a desenhar uma 'curva' de 'tempo de vida média' do *paper*, em que ano/anos atinge o pico de impacto, como decresce, etc.. (*"ageing factor"*)

- RIM: Recent Impact Method
- SIM: Subsequent Impact Method
- PIM: Parallel Impact Method
- *conversion of GF*, etc.

n.b. a questão de novo é que os estudos aqui apresentados se referem quase só à *Natural Sciences (Chemistry)*. Será de presumir que noutras áreas, v.g., A&H, é igual? Não... (p.75)

- *Appropriate for predicting?* O maior impacto parece dar-se entre o 2º e o 4º ano. Entre os 17-20 anos praticamente "morreu"...

VI - The ageing of scientific information

- The ageing is similar in most fields of *Natural Sciences*
- «The ageing of scientific information as tackled above [only] refers to quantitative aspects of the phenomenon. The way of ageing, development, modification, incorporation and institutionalisation of scientific information is, however, a quite complex social process.» (p.78)



**VII – Scientmetric indicators for
assessment of publications**

VII – Scientmetric indicators for assessment of publications

- α) **TTP** - Total Publication Productivity [n.b.: books not included!]
- β) **JPP** - Journal Paper Productivity
- γ) **JPC** - Journal Paper Citedness
- δ) **CPR** - Citation per Researcher
- ε) **AI** - Activity Index
- ζ) **TI** - Atractivity Index
- η) **PS** - Publication Strategy
- θ) **RPS** - Relative Publication Strategy
- ι) **RCR** – Relative Citation Rate
- κ) **RW** – Relative Subfield Citedness
- λ) *π-index, h-index*

VII – Scientmetric indicators for assesement of publications

h-index

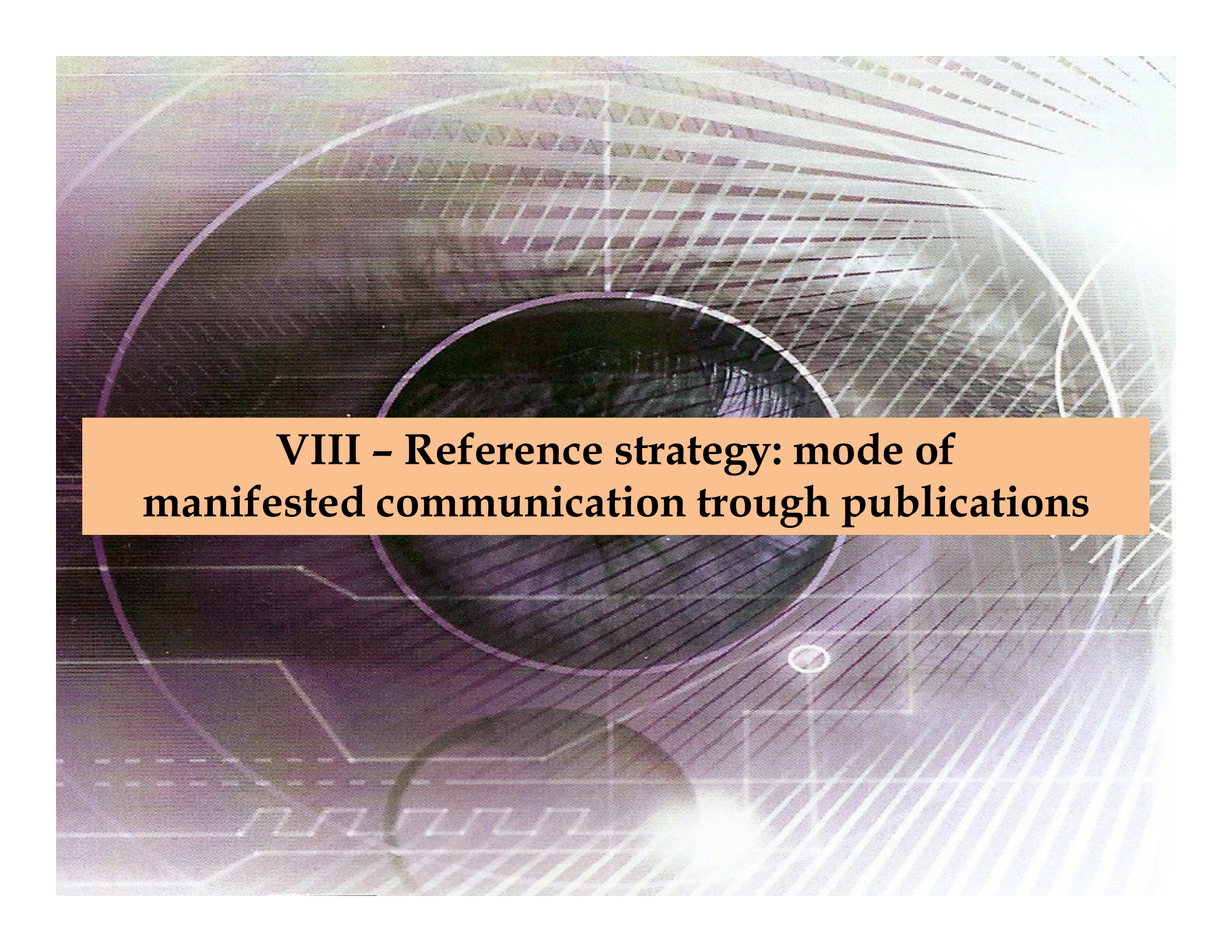
(Hirsch, 2005)

- *h-index* measures the broad impact of an individual's work:
«A scientist has index h if h of his N_p papers have at least h citations each, and the other $(N_p - h)$ papers have fewer than h citations each.»
- V.g.: A. Gradim «tem $h = 110$, se tiver escrito 110 *papers* com pelo menos 110 citações cada um.»

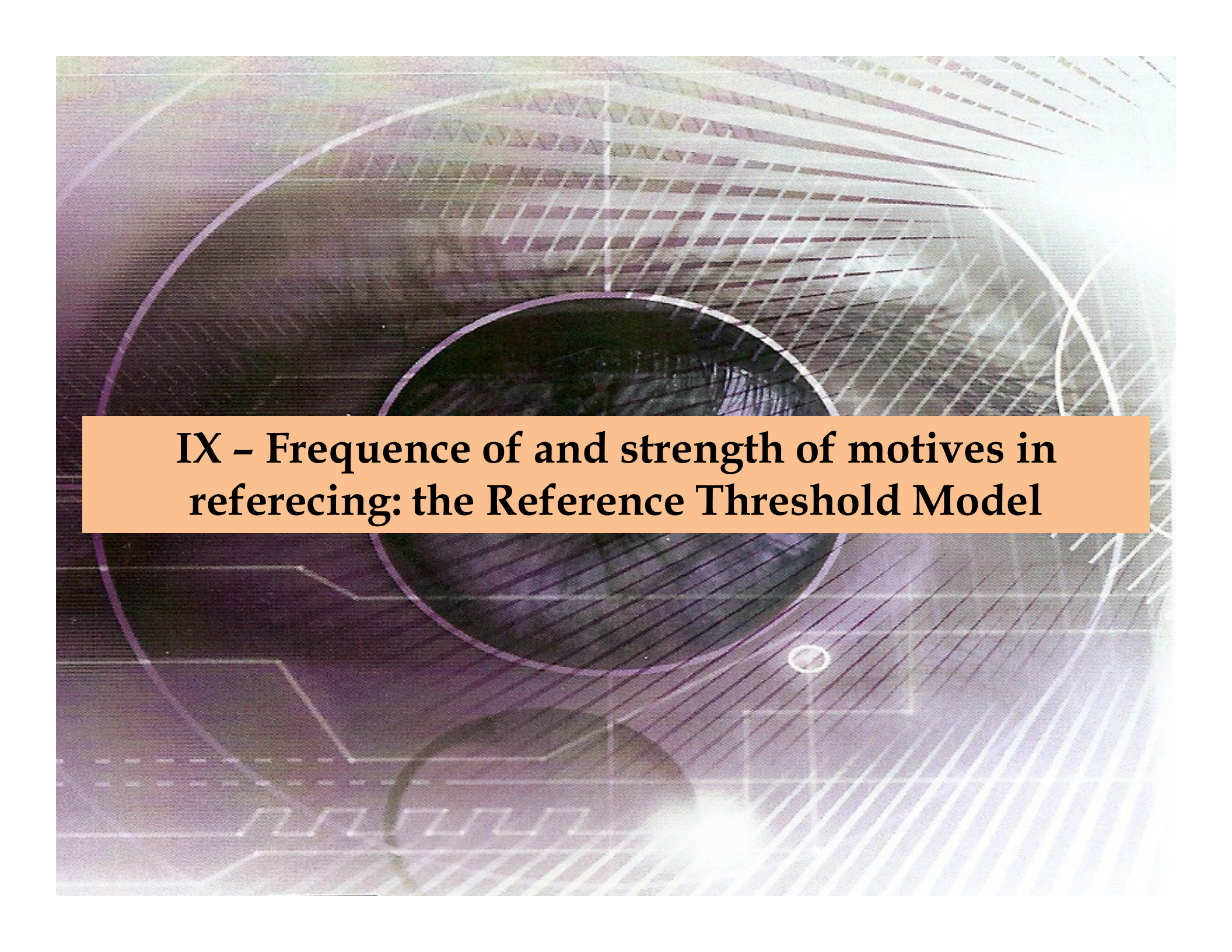
VII – Scientmetric indicators for assessment of publications

Críticas de P. Vinkler ao *h-index*:

- **“Eminence” só pode ser medida pelo “impact” e não pelo “amount”.**
- **O valor do *h-index* só pode ser igual ou menor do que o N_p** (alguém que publicou 25 papers só pode obter $h = 25$, mesmo que os *papers* tenham sido citados 10.000 vezes a revista em que foram publicados tenha, assim, um alto GF)
- **O *h-index* considera igualmente as citações e as auto-citações.** O que dá vantagem a um autor que se cite a si mesmo muitas vezes; enviesa resultados especialmente quando aplicados em ‘low levels’ (I&D, teams, Univs.). Sem auto-citações o *h-index* baixa drasticamente. Também não *normaliza* (‘relativiza’) as citações dependentes (A cita B, porque B cita A, etc.)
- **Os investigadores menos prolíficos são notoriamente prejudicados** (mesmo que um seu artigo tenha sido citado mais vezes, GF), visto que o N_p representa o limite máximo para o *h-index*.
- **O *h-index* não tem em consideração as diferentes características** de cada área e sub-área.



**VIII - Reference strategy: mode of
manifested communication through publications**



**IX - Frequency of and strength of motives in
referencing: the Reference Threshold Model**

IX – Frequency of and strength of motives in referecing: the Reference Threshold Model

— «Considering the role of science as an information-producing system with output exceeding input, we may conclude that the scientific text and references in a publication must be **sufficient** both qualitatively and quantitatively to make the information content of the publication understandable, and the published results reproducible.» (p.137)

— A citação / referência como parte integrante da ciência actual, usada como:

- a) *Links between different topics*
- b) *Development theories and methods*
- c) *Evaluate the impact of particular publications*

IX – Frequency of and strength of motives in referencing: the Reference Threshold Model

— Nas análises cientométricas, os critérios pressupostos para aplicação de referências/citações são:

- a) Que parte significativa do impacto dos autores se manifesta em referências [*“cognitive coupling”*]
- b) Que a maiorias das referências são [*apenas? sobretudo?*] profissionalmente motivadas;
- c) Que têm de ser calculadas independentemente da revista / autor citados

«**Each citation represents a brick in the building of science**» (Price, 1961)

«[Citation] is an ***intertext***, and represents ***the smallest meaningful unit of bibliography.***» (Atkinson, 1984)

IX – Frequency of and strength of motives in referecing: the Reference Threshold Model

McCain and Turner (1989), *et al* - citation context classification scheme:

— “**Utility Index**”:

- a) nº cits.
- b) Onde ocorrem (“the section in wich the referece occurs is a major indication of the use of the referenced paper»): **introduction (15%), methodology (30%), results (30%), discussion and conclusion (25%)**. [Maričić, 1998, encontrou uma correlação forte entre a localização da citação e o seu nível (baixo ou alto)]
- c) Contexto de cada citação (“central or peripheral”)
- d) *Links* explícitos entre *papers* e autocitação (individual e institucional)

IX – Frequency of and strength of motives in referecing: the Reference Threshold Model

«The are numerous papers in the literature for and against the use of citationsfor assesement purposes. Up to now, only a few papers have been published on basis of quantitative data that are against the use of citations for determining the impact of scientific information [MacRoberts, 1989, 1997].»

The nature of the manifested *coupling* between publications referencing and referenced:

- a) Professional (relações científicas e técnicas; *substância*)
- b) Professional and connectional
- c) Connectional (relações sociais entre autores e filiações)

Motivational Strength [M.S.]: absolutely necessary (3), important (2), less important (1). [RTM]


IX – Frequency of and strength of motives in referecing: the Reference Threshold Model

O RTM proposto por Vinkler, tendo em conta razões científicas e relações sociais nas citações, diferencia assim as publicações em:

- a) Professionally relevant (62%)
- b) Relevant for professional and connectional reasons (11%)
- c) Relevant purely for connectional reasons (1.2%)

«The motivational strengths of referencing reported by different authors **are difficult to compare** because of the different terms and types of motivations. Therefore, the value of strongest professional motive in each set given by the corresponding author was chosen to be 100%. The data mentioned support the assumption that **the referencing process is governed primarily by professional** (scientific, informational) **factors**, whereas non-professional reasons play a relatively negligible role.»

N.B. - Mas Vinkler terá em conta o *enviesamento* típico de um de questionário deste género?
[n.b. 'missing' references]



X – Research contribution and share of credit of individual authors

X – Research contribution and share of credit of individual authors

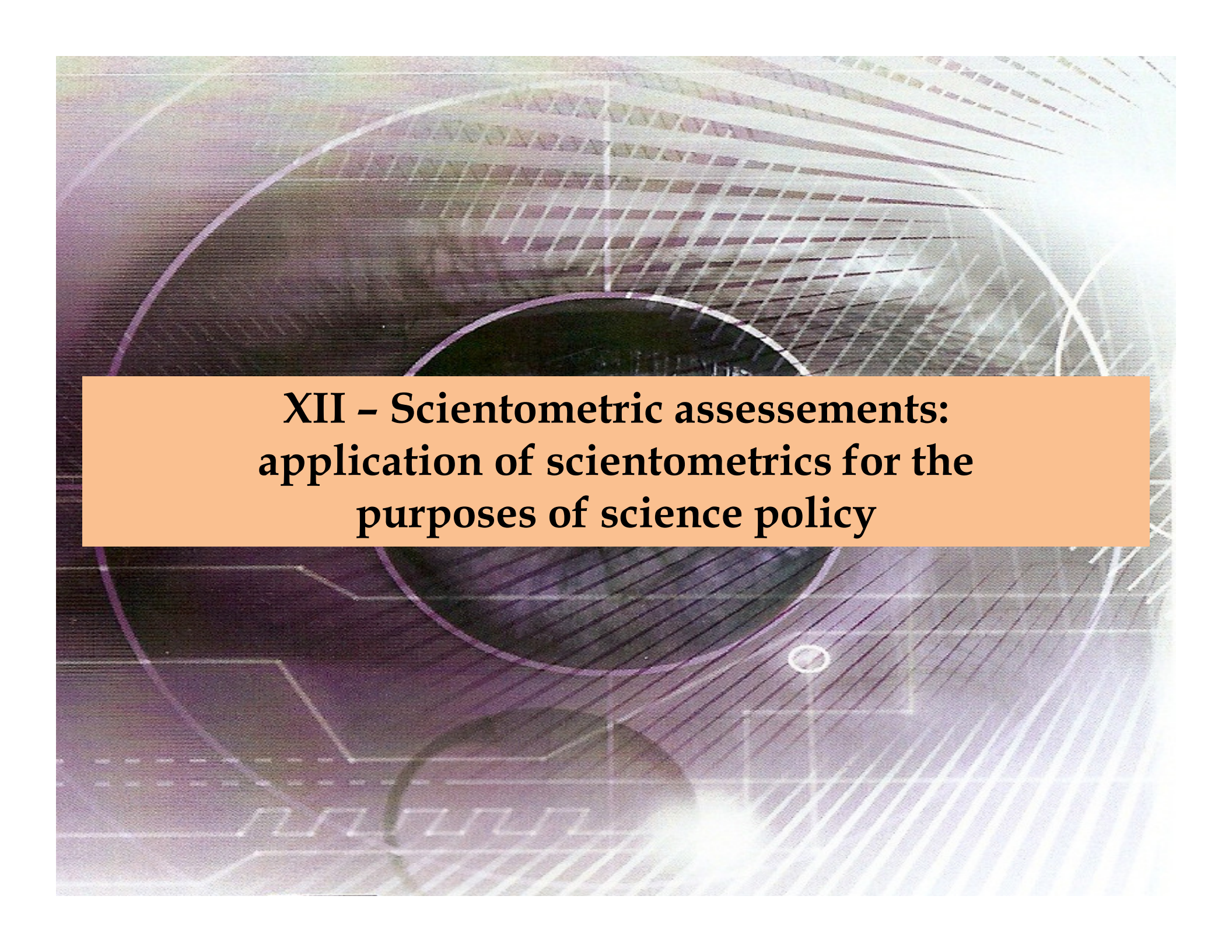
- O lugar essencial da Equipa na ciência actual, quer na teorização (menos) quer na experimentação (mais). A múltipla autoria é assim um fenómeno hoje comum.
- Importância da creditação: para projectos, financiamentos, carreiras, prestígio, etc.
- Como determinar a contribuição de cada co-autor de um *paper*? (55%-60% do total de esforços, num *paper* em Química, vão para a experimentação e análise de dados)

X - Research contribution and share of credit of individual authors


- Dar a cada um dos autores “full credit» (!) (Cronin and Overfelt, 1994)
- “*Fractional counting method*”: a cada um, uma parte (Moed, 2000). V.g., 3 autores: 1/3 para cada um.
- “Distribution of the credit according to the rank of the authors”. “Name-ordering”: fonte de conflitos entre co-autores (Zuckerman, 1968)
- p. 157, Métodos - “geen ketter zonder letter”: Autoria por ordem decrescente de importância; autoria por ordem crescente de importância (Ellwein *et al.*, 1989). Em *papers* com mais de 5 autores, verificou-se que a média mais alta ia sempre para o último nome. (Vinkler, 1993)
- No accepted rules are recognised... [‘connectional’ motives?]



XI - Standards in scientometric assessments



**XII - Scientometric assessments:
application of scientometrics for the
purposes of science policy**



**XIII - Institutionalisation of Scientific Information:
a scientometric model (ISI-S model)**



XIV - Conclusions

XIV – Conclusions

«No *a priori* laws exists!»

- In scientometrics, **'laws' cannot be regarded as strict rules;** they are rather **probabilities** [statistical character].
- **Predictions** are possible? Yes..., a paper with high GF has higher probabilities to be cited than a paper with low GF.
- Nevertheless, one should always take into account the **skewed distribution** on citations...
- **Lack** of generally accepted theories, definitions, rules and classification of knowledge...
- This book is a ***uniform proposal of evaluative scientometrics***

XIV – Conclusions

«No *a priori* laws exists!»

— Dynamic change of information: growth, ageing (6%/y), ... the main research trend can be detected by scientometrics *mapping [in what areas?]*

— Science is an information-producing activity, the essence of which is *communication* [just *journals!*, no *books!*, [p. 83](#)]

— GF [*citation*] may be regarded as ‘ultima ratio’, the “philosopher’s stone in scientometrics”.

— *but* «One of the most serious problems in evaluation of the eminence of journals is how to consider the different effects of bibliometrics factors on the number of citations in different fields.» [“appropriate reference standards should be found!]

XIV – Conclusions

«No *a priori* laws exists!»

- “Mathew effect”: os investigadores com mais alto *standard* tendem a receber mais citações do que merecem. Mas também se observa o inverso: muitos *papers* com alto GF são citados menos que a média **[Invitation Paradox]** (Mt 22, 14)
- Atenção à importância dos “**non-scientific factors**” (‘**connectional**’ or ‘social’) na *journal’s PS*
- Variação diferente dos “relative indicators” em diferentes países (publicação em revistas sem ou com baixo GF, donde que a *RPS* também seja baixa)
- A ameaça para os “evaluative scientometrics” de publicações recentes sem “scientometric knowledge”, sem se considerarem o “*physical meaning of the data*” ... (*pro domo sua?* “retórica cientométrica?”)

XIV – Conclusions

«No *a priori* laws exists!»

- Questões éticas: utilização de Índices “martelados” (“*artefacts*”)
- Muitos indicadores podem ser determinados, mas (na opinião do autor), só alguns poucos podem ser usados para avaliação.
- O autor refere a existência de outros índices “non-scientometrics” (‘social’): *peer review*, PhD numa I&D, nº membros, prémios, etc., mas conclui que “only those indicators that are in accordance with some basic scientometric considerations may be recommended for scientometric evaluation”, e, acima de tudo, “the citation impact” (GF, π -index, [*h-index não*], ...)

XIV - Conclusions

«Peer assessment are limited in time and number of evaluators, and they are planned and orientated.

In contrast, evaluation by citation is unlimited in time and number of evaluators, and the process is not centralised and orientated.»

XIV - Conclusions

«There are several methodological difficulties in applying scientometric relationships and indicators at lower hierarchical levels (e.g. individuals). (...) Indicators calculated for individual ('real') sets, always differ from the indicators calculated for a theoretical ideal 'whole' set.»

(p.265)

XIV - Conclusions

- “Scientometric assessment cannot be performed as a routine.” “It is no possible to prescribe an assesement procedure, wich holds universally.” (p.264)
- «Direct comparion of the performance indicators of teams working in different fields is not possible.» (!)
- Responsabilidade ética e política dos “scientometricians” no fornecimento de dados às instâncias de decisão (*science policy*).

XIV - Conclusions

Assumptions

- the citation is regarded as the scientometric unit of impact of scientific information,
- higher scientific impact is revealed by a higher number of citations,
- the journal paper is regarded as the scientometric unit of information in the fields of natural and life sciences,
- the amount of information produced is related to the number of journal papers published,
- publishing more frequently in periodicals of high eminence means a greater impact may be attained.

XIV – Conclusions

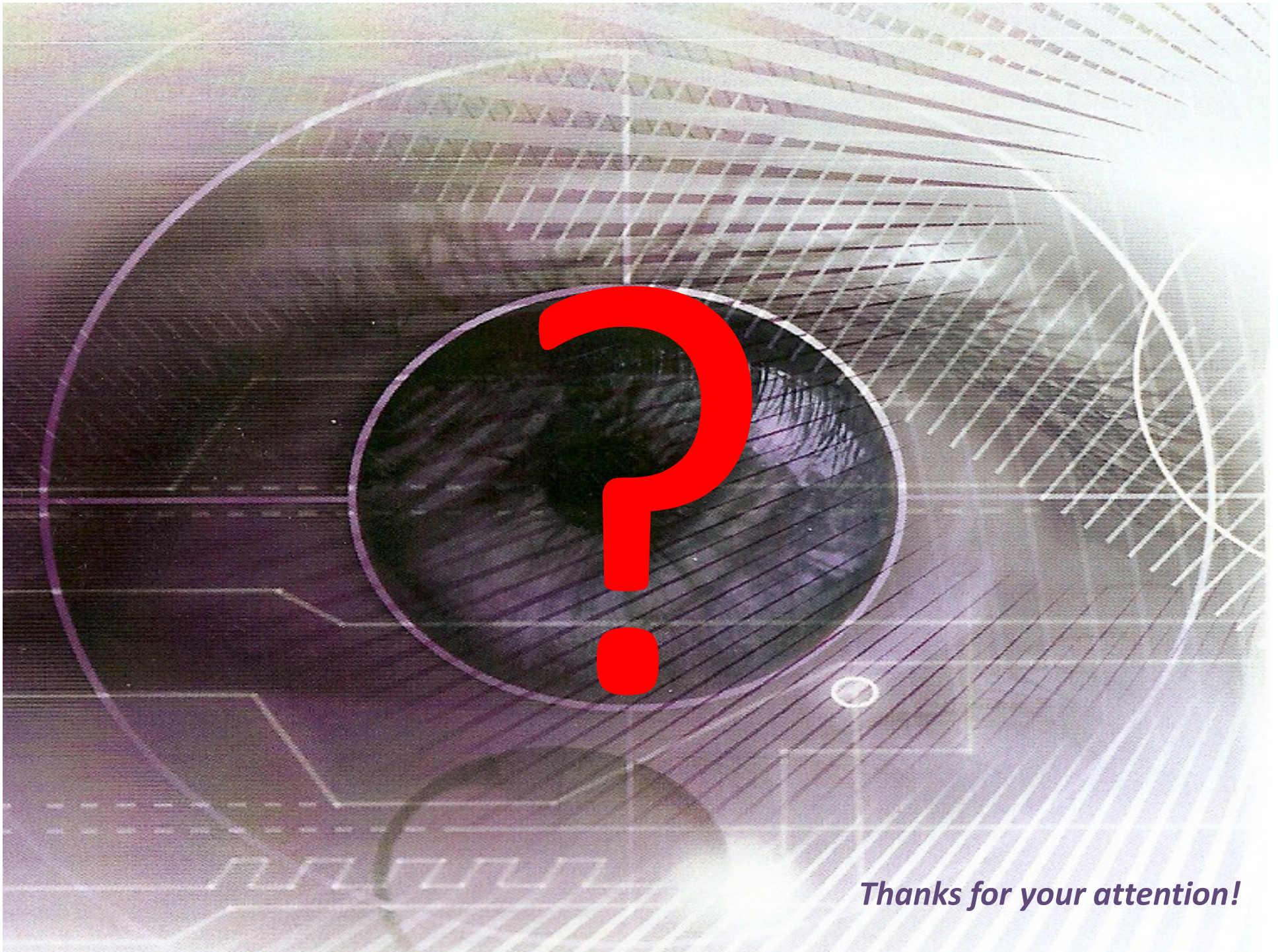
Two opinions

Concerning the evaluation of scientific eminence, one group of scientists is of the opinion that:

- Scientific eminence cannot be or, even should not be, represented by quantitative indicators. Scientific results can be evaluated only by experts in the specific field. Scientometric indicators are not reliable for assessing individuals or teams; only publications and patents, etc., of countries might be analysed by indicators.

The second group of scientists believes that:

- Quantitative indicators may represent scientific eminence ‘more or less’ correctly.



Thanks for your attention!