

**UNIVERSIDADE DE ALCALÁ DE HENARES
UNIVERSIDADE LUSÓFONA DE HUMANIDADES E
TECNOLOGIAS**

**DOUTORAMENTO EM CIENCIAS DOCUMENTALES
BIBLIOGRAFÍA Y DOCUMENTACIÓN**

**Indicadores de visibilidade, impacto e influênciã científica:
a auto citação na análise e avaliação da investigação científica**

Trabalho apresentado por Maria João Amante

**Curso: Utilización de los indicadores bibliométricos para el análisis y
evaluación de la investigación científica
Profª Doutora Mª Ángeles Zulueta Garcia**

1. Introdução

Na avaliação da investigação, os indicadores de citação constituem importantes instrumentos para avaliar o impacto do que se escreve sobre Ciência. Mas a utilização de informação bibliométrica sobre citação tem estado envolvida em grande polémica sobretudo a partir do momento em que os indicadores bibliométricos, para além de serem aplicados com o intuito de monitorar o desempenho institucional ou nacional em termos de investigação, passaram a integrar a fórmula de financiamento da investigação científica. Trata-se então de saber, em que medida as citações reflectem a “qualidade da investigação” e, se assim o entendermos, em que medida esses indicadores são fiáveis dado que os autores podem deliberadamente influenciar ou manipular o impacto mensurável das suas publicações, por exemplo, através do recurso à auto citação. Por este motivo, os utilizadores dos resultados bibliométricos consideram a auto citação como uma forma artificial de inflacionar as taxas de citação e de assim afirmar o posicionamento profissional e de investigação de um autor na comunidade científica. Por este ser um tema sobre o qual estamos longe de alcançar o consenso pareceu-nos interessante aprofundar o seu estudo, partindo de uma breve síntese sobre as origens, evolução e objectivos da Bibliometria, passando pela análise de citação e terminando com a apresentação de alguns estudos e seus resultados sobre a auto citação, desenvolvidos por diferentes autores. Pensamos que este assunto não pode ser dissociado da análise sobre a forma como é desenvolvida e comunicada a actividade de investigação científica, a qual tem crescentemente acentuado a característica de colaboração ao nível regional, nacional e, até transnacional.

2. Bibliometria (origens, breve evolução e objectivos)

O termo *Bibliometria*, introduzido por Pritchard em 1969, em substituição do termo *bibliografía estatística*¹, significava “...el conjunto de estudios que tratan de cuantificar el proceso de la comunicación escrita y la naturaleza y evolución de las disciplinas científicas (como se reflejan en la literatura) mediante el recuento y análisis de diversas características de dicha comunicación.” Podem ser ainda consideradas outras definições, a saber:

- “estudio cuantitativo y análisis de todas las formas de comunicación escrita que adopta la literatura científica”
- “aplicación de métodos matemáticos y estadísticos al estudio de la literatura científica y de los autores que la originan”
- “actividad interdisciplinar que forma parte de la ‘ciencia de la ciencia’”
- “resultado de aplicar el método científico al estudio de los resultados de la ciencia.”²

Simultaneamente, Nalimov e Mulchenko, citados por Glanzel (2003), falavam de *Cientometria* como “...the application of those quantitative methods which are dealing with the analysis of science viewed as an information process”³.

¹ O termo *bibliografía estatística* era usado para descrever o domínio de estudo relativo à aplicação de modelos matemáticos e de estatísticas à investigação e para quantificar o processo de comunicação escrita.

² Zulueta Garcia, M^a Ángeles (2007) – *Utilización de los indicadores bibliométricos para el análisis y evaluación de la investigación científica*. [policopiado].

De acordo com estas definições, a *Cientometria* diz respeito à avaliação da comunicação de ciência enquanto a *Bibliometria* lida com processos de informação mais globais. Nas últimas três décadas, as fronteiras entre os dois termos quase desapareceram.

O domínio da Bibliometria evoluiu ao longo dos anos tendo recebido um grande impulso nos anos 70, altura em que a recolha de dados ainda era feita de forma manual. Nessa altura, a Bibliometria "...was characterised by the personalities of enthusiastic researchers much in the way of a 'hobby' to later integrate interdisciplinary approaches as mathematical and physical models on one side, and sociological and psychological methods on the other, not speaking of the long tradition of library science."⁴

Desde os anos 80, a Bibliometria tornou-se uma disciplina científica distinta com um perfil de investigação específico, vários sub domínios e estruturas de comunicação científica de que são exemplos a revista internacional *Scientometrics*, publicada desde 1979, a realização de conferências internacionais desde 1983, a publicação da revista *Journal Evaluation*, publicada desde 1991 e a criação, em 1995, de uma associação profissional internacional, a *International Society for Scientometrics and Informetrics*.

Para Glanzel (2003) as causas deste desenvolvimento residem na disponibilidade de grandes bases de dados bibliográficos e no rápido desenvolvimento da ciência e da tecnologia de computadores.

Nesta linha de pensamento Zulueta Garcia (2007) considera que o desenvolvimento dos estudos bibliométricos se deve:

- "Razones técnicas, fundamentalmente debidas a la informatización de las bases de datos bibliográficas en todas las áreas, el fácil acceso a las mismas (desarrollo de las telecomunicaciones) y la creciente estandarización que permiten el desarrollo de los nuevos indicadores
- Necesidad de evaluar la actividad científica, debido a los crecientes costes de la investigación y a las restricciones económicas y a la necesidad de evaluar la eficacia de las políticas científicas en cuanto a resultados de la investigación."⁵

Nos nossos dias, a investigação bibliométrica dirige-se a três grupos alvo principais:

- a) Bibliometria para bibliometras (metodologia) sendo aqui que principalmente se realiza a investigação metodológica;
- b) Bibliometria para as disciplinas científicas (informação científica);
- c) Bibliometria para a gestão da política de ciência (política científica).

Constituem objectivos específicos da Bibliometria:

- a) A realização de estudos descritivos que analisam a dimensão, o crescimento, a distribuição e a dispersão dos documentos científicos;
- b) A realização de estudos sociométricos que analisam a estrutura e a dinâmica dos grupos que produzem os documentos e dos consumidores;
- c) O estudo das relações existentes entre os vários elementos bibliográficos contidos nos documentos científicos.⁶

³ Glanzel, W. (2003) – *Bibliometrics as a research field: a course on theory and application of bibliometric indicators. Course handouts*, 6.

⁴ Op. cit., 9.

⁵ Zulueta Garcia, M^a Ángeles (2007) – *Utilización de los indicadores bibliométricos para el análisis y evaluación de la investigación científica*. [policopiado].

⁶ Apesar de, nos nossos dias, a Bibliometria ter estes objectivos específicos, Glanzel (2003) lembra a existência de algumas crenças erróneas quanto à Bibliometria por exemplo o facto de, para alguns, "...the main task of bibliometrics should be the expeditious issuing of 'prompt' and 'comprehensible' indicators for science policy and research management", "research on bibliometric methodology is

Na perspectiva da metodologia, podemos considerar três níveis de agregação na investigação em Bibliometria. Cada nível de agregação requer uma abordagem metodológica e tecnológica específica. Considera-se, assim, a existência do nível micro (relativo ao *output* de publicação de indivíduos e de grupos de investigação), do nível meso relativo ao *output* de publicação de instituições e estudos de revistas científicas) e do nível macro (*output* de publicação de regiões e países e de agregações supranacionais).

De acordo com Glanzel (2003) esta distinção é necessária por várias razões a saber “...the mathematical-statistical background, the precision of retrieval and cleaning-up of data, different counting schemes, different meaning of bibliometric conceptions (e.g. self-citations), non-additivity of bibliometric data (because of multiple assignment).”⁷

Para o desenrolar da sua actividade, a Bibliometria desenvolveu e aplica um conjunto de leis, denominadas leis bibliométricas (Price – crescimento exponencial da Ciência; Lotka – produtividade dos autores; Bradford – dispersão da literatura científica; envelhecimento da Ciência) que se baseiam no comportamento estatístico regular que os vários elementos que integram a Ciência evidenciaram ao longo do tempo. Igualmente foram desenvolvidos indicadores, os indicadores bibliométricos (de produção, de visibilidade, impacto ou influência científica e de relação) os quais constituem uma medida que proporciona informação sobre os resultados da actividade científica em qualquer uma das suas manifestações.

3. Objectivo do estudo e fontes

O nosso estudo incide nos indicadores de visibilidade, impacto ou influência científica e, mais especificamente, na análise de citações centrando-se o tópico na auto-citação. Para o tratamento do tema foi realizada uma pesquisa na *ISI Web of Science* com o objectivo de identificar um conjunto de artigos que abordassem o tema da auto citação quer numa perspectiva teórica quer numa perspectiva aplicada, isto é, artigos que referissem projectos de investigação concretos neste domínio. Assim, foi identificado o artigo *A concise review on the role of author self-citations in information science, bibliometrics and science policy*, da autoria de Glanzel, W., Debackere, B. T. e Schubert, A., publicado em 2006 na revista *Scientometrics* que, conforme referimos anteriormente, constitui uma revista de referência na área da Bibliometria. Com base nas referências bibliográficas apresentadas na bibliografia, foram identificados outros artigos por nós considerados relevantes para o tratamento do tema (devidamente assinalados na Bibliografia). Foram ainda utilizados outros documentos como fontes adicionais de informação, constando igualmente da Bibliografia deste trabalho.

4. A análise de citação

A análise de citação teve início com a publicação do *Science Citation Index* (SCI) em 1961. Desde essa altura, o método cresceu muito em popularidade. Assim, na avaliação

unnecessary, instead bibliometricians should elaborate guidelines explaining the use of their indicators”, “bibliometrics might be reduced to simple counting activities in order to replace/supplement qualitative assessment by quantitative indicators and to set publication output off against funding.”⁷

⁷ Glanzel, W. (2003) – *Bibliometrics as a research field: a course on theory and application of bibliometric indicators. Course handouts.* 37

da investigação os indicadores de citação situam-se entre as mais importantes medidas de impacto da literatura científica. Contudo, a aplicação de dados bibliométricos sobre citação tem-se revelado uma questão controversa.

A própria noção de autoria e a sua atribuição tem colocado questões ao longo dos tempos. No século XVII, antes do aparecimento das revistas académicas, os cientistas comunicavam através de cartas. A troca de correspondência constituía o principal meio para a partilha de ideias e de resultados experimentais. O aparecimento, em França, da revista *Journal des Sçavans* (1665) e, em Inglaterra, da revista *Philosophical Transactions* “...constituted the beginnings of the journal-based scholarly communication system as we know today.”⁸

Nos nossos dias, a questão da autoria ganhou novos contornos dadas as características altamente competitivas do mercado em termos de publicação, no qual a reputação, o sucesso em termos de carreira e até a remuneração estão intimamente ligados à predominância em termos de publicação e de citação.

Segundo Glanzel, Debackere e Schubert (2006) a discussão sobre a utilização de indicadores de citação ganhou uma dimensão nova quando os indicadores bibliométricos, para além de serem usados para avaliar o desempenho da investigação institucional e nacional, passaram a integrar a fórmula de financiamento da investigação científica. Tornou-se assim, imprescindível saber em que medida as citações reflectem a “qualidade da investigação” e, se assim for, saber se essas medidas são fiáveis dado que os próprios autores podem manipular o impacto mensurável das suas publicações. “Thus, users of bibliometric results are sometimes condemning author self-citations as a possible means of artificially inflating citation rates and thus of strengthening the authors’ own position in the scientific community.”⁹

Os indicadores de visibilidade, impacto ou influência científica baseiam-se, entre outros instrumentos, na contabilização do número de citações que recebem os documentos desde o momento da sua publicação. Estima-se que, em média, cada artigo apresente 15 referências bibliográficas.¹⁰ A recepção de citações é muito assimétrica aspecto evidenciado pelo facto de 35% não receber nenhuma citação, 49% receber apenas uma citação e 8%, 2 citações, 3%, 3 citações, 2%, 4 citações e 1%, 5 citações ou mais.

Moravcsik, referido por Zulueta Garcia (2007), classifica as citações de acordo com a sua natureza e função. Assim, podemos falar dos seguintes tipos de citações:

- *Conceptual* – teórica – em que um conceito ou teoria é usado, directa ou indirectamente, no artigo vs *Operativa* – metodológica – quando o conceito se menciona como instrumento para apoiar a afirmação do autor;
- *Orgânica* – essencial – no caso das referências aos conceitos e teorias que servem de base à tese do artigo vs *Superficial* – não essencial – no caso de referências a procedimentos alternativos ou métodos que constituem rotinas na literatura da especialidade, não sendo realmente necessárias para o artigo;
- *Evolutiva* – desenvolvimento de uma ideia – no caso das referências que contribuem para o desenvolvimento lógico do tema do artigo vs *Justaposição* – contraste de uma ideia – no caso de referências a trabalhos similares que

⁸ Cronin, B. (2001) – Hyperauthorship: a post modern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*. Vol. 52 (7), 559. ISSN 1532-2882.

⁹ Glanzel, W., Debackere, K. e Schubert, B. T. (2006) – A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics*. Vol. 67 (2), 263. ISSN 0138-9130.

¹⁰ Destes, 10% não incluem nenhuma referência; 85% tem menos de 25 referências (50% do total); e 5% tem mais de 25 referências (os restantes 50%).

utilizaram procedimentos análogos, mas que não contribuem para o desenvolvimento lógico do tema;

- *Confirmativa* – apoio a um argumento – quando o autor considera que o artigo que cita é correcto vs *Negativa* – oposição ao mesmo – quando o autor cita o artigo mas considera-o errado.

Esta classificação de citações tinha como principais objectivos distinguir entre “...las ideas y los procedimientos usados en el artículo, las citationes necesarias y las prescindibles, material semejante al artículo de aquel que sigue líneas divergentes, entre material considerado bueno y los que son juzgados malos.”¹¹

As razões para a citação podem ser agrupadas em *razões sérias ou funcionais* (influência sobre o conhecimento)¹², *razões estratégicas ou menos sérias* (motivações sociais)¹³ e *motivações políticas ou ideológicas*.¹⁴

Porque, como acabamos de referir, as razões para a citação nem sempre são as mais justificáveis e correctas do ponto de vista ético, MacRoberts e MacRoberts (1988) enunciam os problemas que a sua análise comporta:

- “1. Formal influences not cited.
 - a) Biased citing.
2. Informal influences not cited.
3. Self-citing.
4. Different types of citation.
5. Variations in citation rate related to type of publication, nationality, time period, and size of speciality.
6. Technical limitations of citation indexes and bibliographies.
 - a) Multiple authorship.
 - b) Synonyms.
 - c) Homonyms.
 - d) Clerical errors.
 - e) Coverage of literature.”¹⁵

Zulueta Garcia (2007) resume estes problemas ao afirmar que os mesmos ocorrem “...porque os autores científicos no siempre siguen comportamientos completamente lógicos o consistentes en sus hábitos de citación. A pesar de utilizarse indicadores

¹¹ Zulueta García, M^a Ángeles (2007) – *Utilización de los dos indicadores bibliométricos para el análisis y evaluación de la investigación científica*. [policopiado].

¹² Nestas incluímos dar crédito a um trabalho relacionado, identificar metodologias, equipas, etc., introduzir o tema e situar antecedentes, corrigir trabalhos próprios ou de outros ou criticar trabalhos anteriores, apoiar afirmações, autenticar dados, identificar artigos originais nas quais uma ideia ou conceito foi discutido e disputar prioridade relativamente a uma descoberta.

¹³ Aqui deverão ser contemplados o reconhecimento pelos autores citados, a publicidade a trabalhos anteriores que não foram difundidos ou que não foram citados, a divisão dos resultados da investigação nas mínimas unidades possíveis para se citarem umas às outras, as referências desnecessárias a figuras importantes, as referências para obter benefícios (dinheiro, políticos, etc), as referências seleccionadas de acordo com as preferências editoriais das revistas para onde se envia o artigo a publicar, o intercâmbio de referências com colegas para aumentar mutuamente o número de citações e o não reconhecimento de autores novos.

¹⁴ A utilização de citações é considerada como uma demonstração da influência que um autor recebe de outros autores, motivo pelo qual a análise de citações e de co-citações constitui um instrumento para medir a rede de influência das ideias científicas. Dado que as citações e as ideias num texto não aparecerem isoladas mas combinadas com outras citações e ideias é, assim, possível aceder a um sistema complexo de símbolos que o autor partilha com outros e que está em permanente mutação.

¹⁵ MacRoberts, M. H. e MacRoberts, B. R. (1988) – Problems in citation analysis: a critical review. *Journal of the American Society of Information Science*. Vol. 40 (5), 343. ISSN 0002-8231.

basados en los recuentos de citas, todavía no existe una teoría adecuada sobre este punto, aunque se han hecho algunos estudios empíricos muy útiles.”

Os principais problemas são de dois tipos: técnicos e conceptuais. Quanto aos primeiros, neles podemos incluir os erros tipográficos, as polissemias (autores com variações nas grafias), as homónimas, a necessidade de ter alguns cuidados na contabilização das citações (ser cauteloso na análise de comportamentos de citação individuais, as análises de citações devem recair sobre conjuntos comparáveis, devem ser evitadas as conclusões relativas à qualidade da investigação). Igualmente, estes indicadores devem ser associados a contabilização das publicações.

Quanto aos problemas conceptuais derivam do facto de documentos com dados errados poderem ser muito citados, de alguns tipos de documentos serem mais citados do que outros com a mesma qualidade (monografias vs artigos), do fenómeno de obliteração, da auto citação, do efeito de halo, da variação de hábitos de citação e de publicação de acordo com as diferentes disciplinas, da sobre-representação de artigos de carácter metodológico e das variações em termos de obsolescência ou envelhecimento da literatura científica de acordo com as diferentes disciplinas, motivo pelo qual não são indicadas as comparações entre disciplinas.

4.1 A auto citação

Na literatura sobre Bibliometria está em curso um debate sobre a interpretação e o papel da auto citação no processo de comunicação científica o qual reflecte alguma polarização: uns, condenam a utilização da auto citação por a considerarem uma forma artificial de inflacionar as taxas de citação reforçando, assim, a posição desses investigadores no seu domínio científico; outros, os bibliometras, consideram adequada a aceitação de parte dessas auto citações como elemento do processo de comunicação científica.

Conforme vimos anteriormente, a auto citação aparece incluída nos problemas conceptuais que se colocam à análise de citação. Vários autores têm afirmado que o termo auto citação é usado para referir as várias formas de relação entre os artigos que citam e aqueles que são citados. Fala-se ainda de auto citação de revistas quando um artigo publicado numa revista cita outro artigo publicado na mesma revista. O termo foi igualmente utilizado para descrever a relação entre o artigo que cita e o artigo citado quando ambos têm como origem a mesma instituição. Frequentemente o conceito é usado para descrever a relação existente entre os autores do artigo que cita e os do artigo que é citado.

Lawani (1981) no seu artigo, ocupa-se apenas deste último tipo de auto citação tendo como objectivo a construção de uma grelha de análise que ajude a clarificar as discussões relativas ao tema auto citação, procurando aclarar a natureza heterogénea da auto citação através da apresentação de um esquema para a sua classificação sistemática.

Considera que as discussões que, sobre o assunto, encontramos na literatura, fornecem uma impressão enganadora sobre o processo de auto citação, como se de uma realidade homogénea se tratasse. Em sua opinião, existem duas classes e, em cada classe, são identificáveis quatro categorias. Pedindo emprestada a terminologia da taxonomia, o autor fala de dois géneros de citação, cada um dos quais com quatro espécies. Os dois géneros são as auto citações síncronas e as auto citações diacrónicas. “...author’s

synchronous self-citations are those contained in the citations the author gives, whereas diachronous self-citations are those included in the citations an author receives.”¹⁶

Ilustra a distinção entre os dois géneros de auto citação através da análise dos primeiros quatro artigos publicados no primeiro número dos *Proceedings of the National Academy of Sciences* (USA, 1977). Os dados apurados permitiram ao autor concluir que um artigo (ou um autor) pode ser caracterizado pelas seguintes relações:

- (1) elevada taxa de auto citação síncrona e baixa taxa de auto citação diacrónica;
- (2) baixa taxa de auto citação síncrona e elevada taxa de auto citação diacrónica;
- (3) elevada taxa de auto citação síncrona e elevada taxa de auto citação diacrónica.

Uma outra combinação possível, mas não revelada, seria uma baixa taxa de auto citação síncrona associada a uma baixa taxa de auto citação diacrónica. Considera que cada uma destas combinações pode ter um significado sociológico que deve ser estudado referindo o caso da vaidade. Em sua opinião, as aplicações sociológicas da sua classificação das auto citações necessitam de ser aprofundadas.

Quanto às quatro espécies apresenta-as da seguinte forma:

- Espécie 1: auto citação na qual o nome do primeiro autor do artigo que cita é igualmente o do artigo citado. Esta pode ser considerada a auto citação clássica;
- Espécie 2: aquela em que qualquer um dos co-autores do artigo que cita é o primeiro autor a aparecer no artigo citado;
- Espécie 3: aquela em que o primeiro autor do artigo que cita é co-autor do artigo citado;
- Espécie 4: aquela em que um co-autor do artigo que cita é igualmente co-autor do artigo citado.

Assim, para Lawani (1982) alguns factores podem afectar as várias espécies de auto-citação apresentadas, a saber, “...the extent of multiple authorships in the research area concerned and the relative standings of the collaborating authors.”¹⁷

Para o autor, a classificação por si proposta tem aplicações sociológicas: uma será o estudo da colaboração em investigação, a outra será a clarificação dos aspectos relativos à vaidade tendo como objectivo a sua possível medição.

No seu artigo, **MacRoberts e MacRoberts (1988)** pretendem contribuir para a resolução do debate entre os defensores da análise de citação e os seus críticos, através do estudo dos problemas da teoria normativa da citação¹⁸ de forma a avaliar o trabalho que sobre eles tem sido feito. Mais especificamente, “...we are concerned with problems of the ‘file’ and its developments rather than with problems after the data have been gathered, such as interpretation, theory testing, and statistical handling.”¹⁹

Consideram que a auto citação tem frequentemente sido considerada como um problema potencial, principalmente por a sua utilização parecer excessiva, com entre 10 e 30% do total de citações. Apesar de terem existido várias tentativas para resolver este “problema” (por exemplo, através da eliminação da auto citação da análise de citação)

¹⁶ Lawani, S. M. (1982) – On the heterogeneity and classification of author self-citations. *Journal of the American Society for Information Science*. Vol. 33 (5), 281. ISSN 0002-8231.

¹⁷ Op. cit., 284.

¹⁸ Esta teoria defende que as bibliografias são listas de influências e que os autores citam de forma a conceder crédito sempre que esse crédito é devido. Isto significa que, sempre que um autor usa informação do trabalho de outro autor, citará esse trabalho. Assim, quando um autor utiliza o trabalho de outro autor, seja directamente ou através de fontes secundárias, e tal é evidente no texto, significa que foi influenciado por esse trabalho.

¹⁹ MacRoberts, M. H. and MacRoberts, B. R. (1988) – Problems of citation analysis: a critical review. *Journal of the American Society for Information Science*. Vol. 40 (5), 343. ISSN 0002-8231.

“...there is little justification for any particular strategy since self-citing has not been studied, except for its basic surface features.”²⁰ Referem o estudo realizado por Tagliacozzo²¹ relativo ao âmbito da auto citação, ao período das auto citações, ao carácter repetitivo das auto citações, à relação entre a autoria múltipla e as auto citações, à relação entre a dimensão da bibliografia e as auto citações e à relação entre a produtividade dos autores e as auto citações. Concluíram que poucos artigos não incluem qualquer auto citação, que a distribuição do número de auto citações por artigo varia muito e que os autores citam o seu trabalho mais frequentemente do que o trabalho de qualquer outro autor. Nesta linha, consideram que estes resultados “...neither support nor reject the contention that self-citation by authors of scientific articles is excessive. Only a careful analysis of the text where citations occurred could perhaps provide a basis for judging whether some of the citation were superfluous.”²² Assim, MacRoberts e MacRoberts (1989) consideram que, sem a produção de informação adicional sobre o tema auto citação a possibilidade de a auto citação ser excessiva está por confirmar.

Pichappan e Sarasvady (2002) consideram que a auto citação tem sido analisada essencialmente na perspectiva da avaliação sendo clara a existência de uma tendência para acentuar as suas desvantagens ignorando-se o papel cumulativo do conhecimento o qual, por si só, justifica que a experiência dos autores, os seus trabalhos anteriores, sejam normalmente referidos através da auto citação.

Destacam que, apesar dos muitos estudos empíricos que sobre a citação têm sido realizados os mesmos “...are not just good with respect to author self-citations; the main reason is the ‘social desirability’ of authors. To what extent authors honestly view their own citations?”²³

Trata-se assim de analisar o comportamento de citação dos autores de forma a formular uma explicação teórica para o comportamento de citação de cientistas e académicos. Em conformidade, segundo Pichappan e Sarasvady (2002) os autores citam-se a si próprios por várias razões:

- “...to add further and rejuvenate the earlier work;
- to increase the visibility of earlier work, (for example, when an author cites his/her earlier work published in a less visible journal, in a widely circulated journal, he/she invites attention of the users to the earlier;
- to keep the paper to stand in the literature. Authors prefer to keep the papers as eternal and to move them in citation cycle;
- to alert forthcoming work;
- to market the poorly cited or uncited work;
- to increase the validity of earlier argument by providing further data and evidences to familiarise the methodology or formula, etc., proposed earlier;
- to establish ‘mastery’ in the area;
- to convince peers and users; referencing can increase the persuasiveness of a scientific paper; and

²⁰ MacRoberts, M. H. and MacRoberts, B. R. (1988) – Problems of citation analysis: a critical review. *Journal of the American Society for Information Science*. Vol. 40 (5), 344. ISSN 0002-8231.

²¹ Cf. Tagliacozzo, R. (1977) – Self-citations in Scientific Literature. *Journal of Documentation*. Vol. 33, 251-265. ISSN 0022-0418.

²² Op. cit., 344.

²³ Pichappan, P. and Sarasvady, S. (2002) – The other side of the coin: the intricacies of author self-citations. *Scientometrics*. Vol. 54 (2), 286. ISSN 0138-9130.

- to inflate own citation rate.”²⁴

Igualmente importa considerar que o impacto de um artigo não depende apenas de o autor se citar de forma recorrente. Depende, e muito, do conhecimento sobre a existência do documento e da disponibilidade (acessibilidade) do mesmo sendo que estes factores não interferem no processo de auto citação. Igualmente, o aprofundamento de um assunto por parte de um autor justifica a auto citação dada a existência de uma ligação cognitiva entre o artigo que é citado e o artigo que cita.

Defendem que o grau de auto citação num artigo, por parte de um autor, depende de duas variáveis: o número de artigos que o autor já produziu numa determinada especialidade e em que medida o autor confina o seu interesse a uma especialidade. Falam assim da existência de uma lógica incremental por detrás da auto citação. Referem outros dois factores que, em sua opinião, explicam a extensão da auto citação: a auto inspiração (porque o grau de integração do conhecimento e a sua transmissão são mais rápidos através da auto inspiração) e o ambiente que rodeia o autor (porque o ambiente determina o conteúdo dos pensamentos do autor). Neste sentido, argumentam que a prática de auto citação deveria ser analisada a partir do *background* daquele que cita e do problema escolhido pelo investigador por dele ser dependente. Referem ainda que, as auto citações atingem mais rapidamente o seu “pico” do que as citações externas (feitas por outros) nos anos seguintes à publicação do artigo mas tendendo a cair para metade à medida que alargamos o período de análise. Dado que o período de auto citação é relativamente menor, este aspecto tem uma influência profunda no cálculo do factor de impacto e do índice de vizinhança (*immediacy index*) de uma revista. Por este motivo, argumentam que o factor de impacto deveria ser modificado não considerando a auto citação.

Quanto à dedução da relevância dos artigos com base no facto de serem citados tal constitui um erro porque nem todos os artigos citados são relevantes nem todos os não citados são irrelevantes. Destacam o facto de as citações irrelevantes afectarem a precisão de alguns estudos especialmente aqueles que se centram no número médio de referências, na taxa de citação (quer dos autores quer das revistas), no factor de impacto e no índice de vizinhança (*immediacy index*) das revistas.

Os dados relativos a auto citação em revistas estão disponíveis a partir do *Journal Citation Reports (JCR)* que compila e apresenta dados de gestão relativos a revistas, facilitando o estudo da auto citação em revistas segundo diferentes perspectivas e a compilação de indicadores. Contudo, a inexistência de uma base de dados com informação relativa a auto citação impede, em sua opinião, que este assunto seja devidamente investigado. Desta forma, “Author citation data requires to be compiled to compare, evaluate and for over all management of author citations. The author self-citation rate offers meaningful interpretation relates to the bibliometric author analysis.”²⁵

Referem ainda o estudo realizado por Lawani (1982), e por nós anteriormente referido, no qual são apresentados alguns indicadores de auto citação. Por considerarem que a taxa média de auto citação constitui um indicador potencial que pode servir como instrumento em muitas decisões relativas à política de investigação, propõem que a taxa média de auto citação de um autor nas várias disciplinas e sub disciplinas seja compilada regularmente de forma a ser determinado o seu posicionamento nas especialidades.

²⁴Pichappan, P. e Sarasvady, S. (2002) – The other side of the coin: the intricacies of author self-citations. *Scientometrics*. Vol. 54 (2), 286. ISSN 0138-9130.

²⁵ Op. cit., 289.

A perspectiva defendida por Pichappan e Sarasvady (2002), ao considerarem a auto citação com um acto privado que coloca questões sobre o que, o quando, o onde e o porquê citar, afigura-se-nos muito interessante. Por esta razão, concluem afirmando que “The preceding arguments do not favour to accept author self-citations, but they lead to more penetrating questions demanding an increasingly subtle level of analysis.”²⁶

Aksnes (2003) apresenta os resultados de um estudo que realizou com o objectivo de avaliar o papel da auto citação na produção científica da Noruega entre 1981 e 1996. Foram analisados 46849 artigos. A percentagem de auto citação face ao total de citações representa 36% quando é aplicada uma janela de análise de três anos mas diminui quando a análise incide em períodos mais amplos. Igualmente, a taxa mais elevada de auto citação diz respeito aos artigos menos citados. Verificou existir uma correlação positiva forte entre o número de auto citações e o número de autores dos artigos sendo que apenas uma pequena parte do aumento global nas taxas de citação que se verifica para os artigos com vários autores se deve a auto citações. De igual modo, a taxa de auto citação revela variações significativas entre as diferentes disciplinas científicas submetidas a análise. Na opinião do autor, os resultados apurados são relevantes para a discussão em curso sobre a utilização dos indicadores de citação na avaliação da investigação.

Considera que os estudos empíricos realizados têm como principal falha a pequena dimensão da amostra (o estudo de Tagliacozzo²⁷ incidiu em 180 artigos; o de Bonzi e Snyder²⁸, em 120 artigos). Igualmente importa considerar que o impacto de um artigo não depende apenas de o autor se citar de forma recorrente. Depende, e muito, do conhecimento sobre a existência do documento e da disponibilidade (acessibilidade) do mesmo sendo que estes factores não interferem no processo de auto citação. Igualmente, o aprofundamento de um assunto por parte de um autor justifica a auto citação dada a existência de uma ligação cognitiva entre o artigo que é citado e o artigo que cita.

Reforça a necessidade de informação mais detalhada sobre as taxas de auto citação dado este aspecto ser de particular importância para a avaliação da representatividade dos indicadores de citação como medidas de desempenho.

O estudo realizado pelo autor tinha por objectivo contribuir para o conhecimento sobre a auto citação e, em particular, sobre as taxas de auto citação. Para tal, foi avaliado o elemento auto citação na produção científica e de que modo a sua taxa varia de acordo com parâmetros como o domínio científico, o total de citações e o número de autores. Por último, discute as implicações dos resultados apurados na utilização das citações como indicadores.

A sua análise foi feita a partir da base *ISI – National Citation Report (NCR)* para a Noruega. Foram excluídos da análise os artigos relativos a Ciências Sociais e a Humanidades. A janela de citação aplicada variou entre os cinco e os vinte anos, com um período médio de citação de 11.6 anos. Para cada artigo, foi calculado o número de artigos citantes que representavam auto citações, utilizando o critério de que, pelo menos um autor, (primeiro autor ou co-autor) é também um autor (primeiro autor ou co-autor) do artigo que cita. Esta metodologia, foi aplicada em várias análises bibliométricas.

²⁶ Pichappan, P. e Sarasvady, S. (2002) – The other side of the coin: the intricacies of author self-citations. *Scientometrics*. Vol. 54 (2), 289. ISSN 0138-9130. .

²⁷ Para mais informação consultar Tagliacozzo, R. (1977) – Self-citations in Scientific Literature. *Journal of Documentation*. Vol. 33, 251-265. ISSN 0022-0418.

²⁸ Para mais informação consultar Bonzi, S e Snyder, H. W. (1991) – Motivations for citations: a comparison of self-citation and citation to others. *Scientometrics*. Vol. 2, 245-254. ISSN 0138-9130.

No total de artigos incluídos na análise (46849), 4845 não foram citados. Dos restantes 42004 artigos, 71% (29842) tinham uma ou mais auto citações. Desde o ano de publicação até 2000, os artigos tinham recebido um total de 640710 citações (média de 13.7 citações por artigo) das quais 136316 eram auto citações representando, assim, a quota de auto citação 21%.

Foi ainda possível apurar que o número de auto citações aumenta com o número total de citações (artigos citados, em média, cinco vezes, receberam 1.4 auto citações enquanto artigos citados, em média, cinquenta vezes, receberam 11 auto citações) e que, em termos relativos, a taxa mais elevada de auto citações se encontra entre os artigos menos citados (esta taxa é de 29,9% para os artigos citados menos de cinco vezes e de 19,4% para os artigos citados de 46 a 50 vezes, baixando a taxa de auto citação para 14,5% nos artigos citados mais de 50 vezes).

Apresenta duas explicações para o facto de um artigo poder não apresentar nenhuma auto citação: o autor não ter continuado a investigar naquele domínio ou, mais frequentemente, o artigo anterior não ter representado uma contribuição relevante para as actividades de investigação do autor posteriormente desenvolvidas.

Assim, a sua primeira conclusão é que, em termos relativos, os artigos com poucas citações apresentam uma taxa mais elevada de auto citação que vai diminuindo à medida que aumenta o número total de citações.

Conforme referimos anteriormente, Aksnes (2003) analisou de que forma o número de auto citações varia em função do número de autores dos artigos. Em sua opinião, seria de esperar que quanto mais autores um artigo tiver maior seja o número de auto citações. Refere a existência de estudos anteriores em que foi provada esta correlação (por exemplo, o estudo de Bonzi²⁹). Refere que, enquanto um artigo com apenas um autor recebe, em média 1.5 auto citações, artigos com dez autores recebem 6.7 auto citações.

Refere ainda que está aceite a ideia de que o número total de citações também aumenta em função do número de autores. Trata-se, assim, de saber em que medida esse aumento pode ser explicado pelas auto citações. Para responder a esta questão, apurou o número global de citações tendo verificado a existência de uma correlação positiva entre o número de autores de um artigo e o número de auto citações: cada autor a mais num artigo traduz-se, em média, em mais duas citações. Contudo, considera que o maior número de auto citações apenas em parte explica o aumento global do número de citações. Em conformidade com este raciocínio, apresenta a sua segunda conclusão: artigos com mais de um autor recebem mais auto citações.

Relativamente à possibilidade de o número de auto citações variar segundo o domínio científico, o estudo realizado confirma essa possibilidade não sendo, contudo, possível, estabelecer correlação quer entre a taxa de auto citação e a taxa média de citação nos domínios em apreço quer entre a taxa de citação e o número médio de autores por artigo. Destaca que, ao nível de cada domínio científico, podem existir outros factores que podem influenciar as taxas de auto citação, a saber, a variação nas normas de citação em cada domínio e a dimensão do trabalho cumulativo ao nível da investigação individual. As percentagens mais elevadas encontram-se ao nível das ciências “duras” o que se explica pelo facto de o conhecimento e o trabalho posterior assentar frequentemente em investigação anterior. Assim, a sua terceira conclusão é que a quota de auto citação varia segundo os diferentes domínio científicos.

²⁹Para mais informação, consultar Bonzi, S. e Snyder H. W. (1990) – Patterns of self-citation across fields of inquiry. *Proceedings of the ASIS annual meeting*. 27, 204-207 e Snyder, H. e Bonzi, S. (1998) – Patterns of self-citation across disciplines (1980-1989). *Journal of Information Science*. Vol. 24, 432-435. ISSN 0165-5515.

Por último, Aksnes (2003) analisa de que forma o número e taxa de auto citação varia com o tempo após a publicação. Para cada artigo foram contabilizadas as auto citações por ano de citação tendo sido calculada a percentagem de auto citação para cada ano de citação. Apurou que a maioria das citações que os artigos obtêm no ano seguinte à sua publicação corresponde a auto citações (63%). Também no segundo ano após a publicação grande parte das citações corresponde a auto citações (40%). Em contraste, quando a janela de análise é ampliada para um período de quinze anos após a publicação, apenas 9% das citações corresponde a auto citações.

Refere ainda que frequentemente, nas análises bibliométricas, as citações são analisadas nos três anos seguintes à publicação do documento. No conjunto da produção científica norueguesa em análise, tal representaria numa taxa de auto citação de 36%. Mas se, em vez de uma janela de três anos, a análise for ampliada para cinco anos, a taxa de auto citação desce para 29%. Conclui então que quando usamos uma janela de análise mais pequena a auto citação revela-se um problema maior do que quando ampliamos o período da análise. Desta forma, o “factor impacto das revistas” será fortemente influenciado pelas auto citações.

Considera a auto citação como um fenómeno ambíguo pois, por um lado, é natural que um autor procure, através da auto citação, estabelecer a relação entre o trabalho em curso e trabalho anterior, mas, por outro lado, pode ser uma manifestação de vaidade. Neste sentido “...authors may tend to cite their own works in order to raise their citation counts or to make their former works visible – although there are practical (frequency of publication) as well as normative limits for how often one can site oneself.”³⁰

Através do seu estudo, conclui que as taxas de auto citação encontradas são mais elevadas do que os resultados apresentados por estudos realizados anteriormente como, por exemplo, o estudo realizado por Lawani (1981), já referido. Tal pode ser explicado pela diferença de metodologias utilizadas já que, no primeiro caso foi feita uma análise diacrónica (na qual a opção por uma determinada janela de citação influencia as taxas de auto citação) e, no segundo, uma análise síncrona.

O autor destaca que o seu estudo é de âmbito nacional podendo os resultados apresentados não ter validade universal mas, por ter que envolvido mais de 45 mil artigos os padrões identificados podem, em sua opinião, ser considerados típicos. Fundamenta esta conclusão no facto de a ciência norueguesa ter uma forte orientação internacional (29% das publicações analisadas têm autores de outros países) que deriva do carácter crescentemente colaborativo da forma como se faz e se transmite Ciência.

Quando se pretendem utilizar a os indicadores de citação como instrumentos da política de Ciência a auto citação pode, em sua opinião, constituir um problema. O facto de o estudo provar que 20 a 35% de todas as citações correspondem a auto citações prejudica a utilização das citações como indicador de impacto científico das publicações. Este aspecto tem sido utilizado, de forma recorrente, pelos críticos da utilização dos indicadores de citação. Refere ainda que, “Although we agree that self-citation is not a major problem when comparing nations (at overall and field levels), this conclusion is more dubious when it comes to universities. The reason is that universities are rather heterogeneous in their research profiles.”³¹

Assim, é aos níveis mais baixos de análise (ao nível micro) que a auto citação representa um problema mais sério, por exemplo, quando avaliamos indivíduos e grupos de investigação. Neste sentido, o problema desapareceria se a auto citação fosse eliminada antes de serem estabelecidas as comparações. Contudo, a eliminação da auto citação das

³⁰ Aksnes, D. W. (2003) – A macro study of self-citation. *Scientometrics*. Vol. 56 (2), 243. ISSN 0138-9130.

³¹ Op. cit., 244.

publicações individuais diminuiria a possibilidade de serem produzidos alguns tipos de indicadores (por exemplo, comparações com taxas médias de citação por domínio). Por outro lado, como foi já mencionado, o número de autores de um artigo tem influência na taxa de auto citação. Os dados apresentados no estudo, mostram que o número médio de autores por artigo passou de 2.6 em 1981 para 4.9 em 1996. Assim, para evitar uma inflação artificial do número de auto citações, o autor propõe a adopção de uma definição mais restritiva para o conceito de auto citação limitando-o, por exemplo, à contabilização das auto citações feitas pelo primeiro autor da publicação. Sugere ainda que, para diminuir o efeito das auto citações, sejam utilizadas janelas de análise mais amplas (sempre superiores a três anos) dado tal contribuir para diminuir a sua quota no total de citações. Igualmente, sugere a realização de investigação futura para comprovar a sua tese de que a níveis mais elevados de agregação (nível macro) os efeitos da auto citação não constituem um problema. Dado que, ao nível micro, conforme foi já referido, as auto citações podem constituir um problema, sugere que sejam eliminadas antes de serem feitas comparações ou que, pelo menos, os seus efeitos sejam considerados antes de as citações serem usadas como indicadores de impacto científico.

Glanzel, Thijs e Schlemmer (2004) analisam o papel da auto citação tentando encontrar algumas regularidades no processo de comunicação científica documentada de forma a lançar as bases metodológicas para um trabalho de campo relativo a uma análise crítica dos padrões de auto citação nos estudos empíricos em qualquer nível de agregação.

O estudo desenvolvido está dividido em três partes: "...the first part of the study is concerned with the comparative analysis of the ageing of self-citations and of non self-citations, in the second part the possible interdependence between self-citations and foreign citations is analysed and in the third part the interrelation of the share of self-citations in all citations with other citation-based indicators is studied."³²

Os resultados do estudo baseiam-se em informação bibliográfica extraída da *Web of Science* (1992-2001) e processada de acordo com indicadores bibliométricos. Os autores referem ainda que, para além da discussão em curso sobre a interpretação e o papel da auto citação, também não existe consenso sobre a forma operativa de definir o conceito auto citação.

Existem, assim, duas abordagens diferentes: ao nível micro (dos autores individuais) uma auto citação directa de um autor A ocorre sempre que A é também co-autor de um artigo que cita um artigo da autoria de A. Esta definição não pode ser aplicada a níveis mais elevados de agregação, isto é, quando as publicações e as citações são agregadas em conjuntos de diferentes co-autores e a noção de auto citação aparece dissociada de uma autor individual A; ao nível meso e macro, devem ser usados outros critérios para determinar o que deve ser considerado auto citação.

Estes autores defendem que "...the large-scale analysis of author self-citations gives interesting insight into the mechanism of scientific communication. Although self-citations indicators are somewhat biased by errors of author identification, self-citation based indicators are valuable supplementary measures that can be used both in informetrics and research evaluation. Because of the already mentioned restriction concerning their reliability, self-citation indicators should be used in addition to traditional citation indicators, but not replace them."³³

³² Glanzel, W., Thijs, B. e Schlemmer, B. (2004) A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*. Vol. 59 (1), 63. ISSN 0138-9130.

³³Op. cit., 76.

Quanto à primeira parte do estudo, os autores consideram que, para um determinado número de citações, existe uma expectativa condicionada de auto citações que pode ser expressa numa lei de raiz quadrada. Este aspecto é, em sua opinião, particularmente relevante por significar que não existe nada de arbitrário no processo de auto citação.

Quanto à hipótese de existência de uma eventual interdependência entre as auto citações e as citações externas, de forma a proceder à sua verificação, foi realizada uma análise de regressão linear. Verificaram que, aplicando uma janela de análise de citação de três anos, a auto citação e a citação externa não são variáveis independentes. A influência e o peso das auto citações no conjunto das citações diminui rapidamente estagnando no terceiro ano após a publicação do documento. Consideram igualmente que a utilização de uma janela de análise de citação não inferior a três anos e não superior a quatro anos é suficiente para uma análise bibliométrica fiável dado a taxa de auto citações face ao total de citações se enquadrar em limites entendidos como aceitáveis pelos autores (cerca de 25% para as ciências da vida; 30 a 40% na ciências naturais e engenharia). Igualmente, a relação entre citações externas e auto citações é dependente do domínio científico em apreço mas revelou ser bastante estável no caso de o ano de publicação e a apropriada janela de análise serem alterados por um período considerável.

Para analisar os padrões de auto citação nacionais, os autores fizeram incidir a sua análise em revistas publicadas no domínio das ciências em 1999. As citações foram contabilizadas através da aplicação de uma janela de auto citação de três anos (1999-2001). Foram igualmente usados dos indicadores nacionais normalizados: o *Mean Expected Citation Rate (MECR)*³⁴ e o *Relative Citation Rate (RCR)*³⁵

Concluem que os autores que, em média, publicam em revistas com factor de impacto baixo, são mais frequentemente citados por si próprios do que por outros o que significa que uma fraca visibilidade aumenta a probabilidade de auto citação. Trata-se assim, da constatação de uma relação entre as taxas de auto citação e os indicadores bibliométricos normalizados (como é o caso do factor de impacto). Concluem afirmando que “The rules derived from the analysis have several implications for research evaluation; they can (...) be used to develop field-specific expected self-citation rates and shares within the framework of the evaluation of research performance in research groups and institutions.”³⁶

A questão da co-autoria, e em que medida esta pode inflacionar o número de auto citações, foi analisada num artigo da autoria de **Glanzel e de Thijs**, publicado em 2004³⁷. Referem o facto de, anteriormente, terem verificado algumas regularidades das auto citações, a saber, o período de auto citação, a relação entre as auto citações e as citações feitas por terceiros e a interdependência entre as auto citações e outros indicadores bibliométricos.³⁸

Consideram existir uma correlação positiva entre o número de autores de um artigo e o número de citações que o artigo recebe num determinado período, verificando-se o

³⁴ Taxa média de citação esperada – é um indicador baseado em revistas que expressa a taxa de citação esperada para um determinado artigo.

³⁵ Taxa de citação relativa – compara a taxa de citação verificada com a esperada nas mesmas janelas de citação.

³⁶*Glanzel, W., Thijs, B. e Schlemmer, B. (2004) A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*. Vol. 59 (1), 76. ISSN 0138-9130.

³⁷*Glanzel, W. e Thijs, B. (2004) – Does co-authorship inflate the share of self-citations. *Scientometrics*. Vol. 61 (3), 395-404. ISSN 0138-9130.

³⁸ Esta análise foi feita no artigo de Glanzel, W., Thijs, B. e Schlemmer, B. (2004) – A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*. Vol. 59 (1), 63-77. ISSN 0138-9130.

mesmo ao nível da cooperação entre indivíduos, entre instituições e entre países. A autoria múltipla acompanha o aumento das auto citações mas o aumento é muito mais fraco do que o verificado no tocante às citações feitas por terceiros. Assim, a autoria múltipla tem como consequência directa o aumento das citações feitas por pessoas sem responsabilidade na autoria do documento. Igualmente, a taxa de auto citações de artigos com apenas um autor é consideravelmente mais baixa do que a verificada para os artigos com vários autores. O número de co-autores no conjunto dos artigos com autoria múltipla não tem, contudo, grande influência na taxa de auto citações (face ao número total de citações) que um artigo recebe. Assim, em sua opinião, os resultados do seu estudo negam a existência de interacção entre o número crescente de co-autores e o número crescente de auto citações (como consequência do maior um maior número de autores envolvidos).

Apresentam as seguintes razões com o objectivo de explicar os resultados a que chegaram:

- A produtividade dos autores envolvidos em artigos com autoria múltipla pode não acompanhar o número de artigos que, em média, os cita;
- Pode existir uma estrutura hierárquica de subordinação entre as equipas de co-autores;
- A autoria múltipla não é o resultado do trabalho de equipas estáveis, isto é, a autoria múltipla pode ter origem numa colaboração ocasional de um ou mais co-autores que, entretanto, não continuam a trabalhar na principal linha de investigação da equipa.

Referem ainda que consideram evidente o facto de os artigos com apenas um autor apresentarem taxas subproporcionais de auto citação.

Num artigo publicado em 2006, **Glanzel, Debackere, Thijs e Schubert** defendem a tese de que para compreendermos o papel da auto citação temos de tomar em consideração dois aspectos: por um lado, a auto citação enquadrada no comportamento de citação existente na comunicação científica e, por outro lado, o papel das citações em geral segundo a forma como são encaradas na Ciência da Informação. Os autores consideram igualmente que, do ponto de vista da Bibliometria, foram já estudadas as regularidades básicas da auto-citação em diferentes níveis de agregação, sendo referidos os estudos de Aksnes (2003)³⁹ e de Glanzel (2004)⁴⁰ já por nós introduzidos neste trabalho. Nesta linha de pensamento, defendem que “Regularities related to the ageing, to the relation between self-citations and foreign citations and to the interdependence of self-citations with other bibliometric indicators have been found which allow the conclusion that at high (macro-) levels of aggregation self-citations can be considered a natural part of scientific communication.”⁴¹ Mas a utilização destes indicadores na avaliação e nas fórmulas de financiamento da investigação apresenta problemas e tem sido entendida como fonte de distorção.

O estudo que desenvolveram, e que é referido neste artigo, teve dois objectivos: por um lado, apresentar os objectivos e as formas da interpretação quantitativa das características bibliográficas na Bibliometria e a sua reinterpretação na política de

³⁹ Aksnes, D. W. (2003) – A macro study of self-citation. *Scientometrics*. Vol. 56 (2), 235-246. ISSN 0138-9130.

⁴⁰ Glanzel, W., Thijs, B. e Schlemmer, B. (2004) A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*. Vol. 59 (1), 63-77. ISSN 0138-9130.

⁴¹ Glanzel, W., Debackere, K. e Schubert, B. T. (2006) – A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics*. Vol. 67 (2), 263. ISSN 0138-9130.

investigação e, por outro, resumir o estado da arte na investigação sobre auto citação em Bibliometria. Consideram que a compreensão do papel das auto citações depende de as considerarmos como parte integrante do comportamento de citação de um autor no quadro da comunicação científica documentada. Analisam, então, a citação na perspectiva da Ciência da Informação, da Bibliometria e da política de Ciência.

Quanto ao primeiro aspecto, começam por considerar que o comportamento de citação é um fenómeno complexo estudado quer pela Ciência da Informação (enquanto meio de disseminação e identificação da informação) quer pela Sociologia da Ciência (que estuda a componente social da informação científica existente na citação). Referem as definições apresentadas por vários autores⁴² para o termo citação considerando que as explicações por eles apresentadas para que sejam citados os trabalhos de outros se podem aplicar às citações que um autor faz a trabalho seu anterior.

Defendem a subdivisão das razões para a citação em três grupos segundo a receptividade de terceiros ao trabalho em causa e a relevância das citações. Quanto ao primeiro aspecto, podemos encontrar citações que expressam uma receptividade positiva, citações neutras e citações que expressam uma receptividade negativa. Quanto ao segundo aspecto, as citações podem ser consideradas relevantes, menos relevantes ou até redundantes. Concluem, assim, que nem todas as citações têm o mesmo peso. Chamam, contudo, a atenção para o facto de a Bibliometria se basear na frequência de distribuição e em funções estatísticas sendo o contexto concreto da citação substituído por indicadores bibliométricos.

Na perspectiva da Bibliometria, os bibliometras começaram por abordar as citações de forma semelhante à dos cientistas da informação, enquanto expressão de algo e enquanto medida da utilização da informação. Assim, os primeiros estudos sobre citações foram feitos com o objectivo de utilizar as medidas de citação nas Bibliotecas. Referem a evolução ocorrida no indicador “factor de impacto das revistas” (*Journal Impact Factor*) nas últimas três décadas: “...the utilisation measure originally designed for the use in libraries, information retrieval and scientific information has evolved to a ‘quality’ measure in the sense of research evaluation for systematic application to science policy.”⁴³

Consideram que as citações podem ser usadas para, de forma directa ou indirecta, avaliar o impacto da literatura científica no quadro da comunicação científica documentada. Desta forma, em sua opinião as auto citações são parte integrante do sistema de comunicação de Ciência dado indicarem uma actividade dinâmica e bem sucedida de publicação em revistas com *referees*. Assim, “Only the almost absolute lack of self-citations over a long period as well as a always overwhelming share of self-citations must be considered ‘pathological’ case.”⁴⁴

A Bibliometria encara as citações como uma medida da receptividade aos resultados científicos. A política de ciência e gestão da investigação encara as citações como uma expressão de impacto ou até de qualidade. Para além deste aspecto, a citação constitui uma forma de recompensa, motivo pelo qual, os críticos da utilização da auto citação como indicador de visibilidade (e de impacto) entendem que esta distorce o sistema ao falsificarem o impacto da investigação. Este aspecto pode levar os autores a evitar

⁴² A este propósito veja-se, por exemplo, o artigo de Pichappan, P. e Sarasvady, S. (2002) – The other side of the coin: the intricacies of author self-citations. *Scientometrics*. Vol. 54 (2), 285-290. ISSN 0138-9130.

⁴³ Glanzel, W., Debackere, K. e Schubert, B. T. (2006) – A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics*. Vol. 67 (2), 266. ISSN 0138-9130.

⁴⁴ Op. cit., 268.

citarem trabalhos seus o que também distorce um comportamento natural de comunicação.

Pelo exposto, em sua opinião, “Consistent policy use of bibliometric indicators might potentially induce changes in the publication, citation and collaboration behaviour of scientists (both positive and negative). If bibliometric tools have an effect on decision-making in science policy and research management and the scientific community recognises the feedback in terms of their funding, then there might be measurable repercussions on their behaviour, too. Re-interpretation and ‘perspective shift’ as mentioned above might even catalyse this process.”⁴⁵

Referindo, entre outros, os resultados de estudos apresentados neste trabalho (Lawani, 1981; Aksnes, 2003; Pichappan e Sarasvady, 2002; Glanzel, Thijs. e Schlemmer, 2004) afirmam que os modelos “puramente aritméticos” tais como o aumento do número de auto citações em função do número de co-autores, ou os modelos “puramente psicológicos” como, por exemplo, que a auto citação está apenas sujeita à vontade e à arbitrariedade dos autores, não podem ser aplicados.

Consideram igualmente que, apesar de os comportamentos individuais de citação poderem apresentar alguns desvios relativamente ao padrão, o facto é que os estudos empíricos realizados não revelaram situações que pudessem ser consideradas problemáticas. Sobretudo ao nível macro, foi confirmado que não existe qualquer justificação para não contabilizar as auto citações no conjunto das citações. Igualmente, as auto citações perdem peso com o passar do tempo “...even more rapidly if self-citations are fractionated by co-authors involved in both cited and citing papers.”⁴⁶

Recomendam a utilização dos indicadores suplementares baseados nas auto citações para compreensão dos padrões de comunicação sobretudo, nos estudos ao nível meso e micro, pois estes indicadores podem ajudar a clarificar se o impacto de citação verificado reflecte realmente a receptividade da comunidade científica aos resultados de investigação apresentados.

5. Considerações finais

Os indicadores bibliométricos têm sido usados recorrentemente para avaliar o desempenho da actividade de investigação. Esta utilização tem por base o princípio de que o objectivo da investigação consiste na produção de conhecimento novo o qual tem como evidência material a publicação, por exemplo, num artigo. O impacto do conhecimento produzido e publicado é medido pelo número de citações feitas a um determinado artigo. Contudo, não podemos entender impacto como sinónimo de qualidade. Por esta razão, a avaliação da qualidade potencial dos artigos deve ser feita tomando igualmente em consideração o prestígio e o nível de impacto das revistas onde são publicados esses artigos.

Da leitura que fizemos dos artigos referidos neste estudo, pudemos concluir que a utilização dos indicadores bibliométricos de visibilidade, impacto ou influência científica é muito mais fiável ao nível macro e meso (em níveis mais elevados de agregação) do que ao nível micro, no qual se avalia o desempenho dos investigadores a título individual. Por esta razão, a validade da aplicação de indicadores bibliométricos para avaliar o desempenho dos investigadores em termos individuais, tem sido alvo de muitas críticas.

⁴⁵ Op. cit, 268-269.

⁴⁶ Op. cit., 275.

Por outro lado, a utilização de indicadores bibliométricos não substitui a avaliação pelos pares (*peer review*). Considera-se assim, que a utilização de dados quantitativos (fornecidos pelos indicadores bibliométricos) deve ser acompanhada por procedimentos de avaliação qualitativos.

Apesar das críticas, a utilização das citações como indicadores de visibilidade, impacto ou influência científica tem-se revelado útil para a avaliação dos resultados científicos e seu impacto em termos de investigação científica, sobretudo a níveis de agregação mais elevados. O número de citações atribuídas pelos autores a um artigo anterior mede o impacto deste artigo num determinado domínio científico. Também é revelador do reconhecimento atribuído ao autor pelos seus pares. Igualmente o número de publicações e de citações que um autor recebe está relacionado com a sua visibilidade enquanto investigador. Contudo, não podemos esquecer o facto de que nem todo o trabalho científico é igualmente visível dado o nível de visibilidade depender, em grande medida, do local e da língua de publicação bem como do domínio científico em apreço. Trabalho de investigação que não tenha visibilidade internacional muito remotamente será citado por outros cientistas para além daqueles que desenvolvem relações de trabalho estreitas com o autor em causa. Daí a relevância de aspectos como o conhecimento sobre a existência do documento e o grau de acessibilidade ao mesmo, os quais têm repercussões nos níveis de citação

Igualmente o recurso à utilização de outros indicadores bibliométricos como é o caso o factor de impacto das revistas ganhou relevância no trabalho científico como instrumento de apoio à gestão da política de investigação.

A evolução verificada na utilização dos indicadores bibliométricos não foi feita de forma pacífica como pudemos confirmar através da leitura dos artigos já apresentados. De domínio de conhecimento de bibliotecários, sociólogos e historiadores de ciência, a Bibliometria passou a ser encarada como alternativa relativamente rápida, fácil e pouco dispendiosa à revisão pelos pares (*peer review*) para avaliação do desempenho da investigação até porque também essa revelou ter algumas limitações.⁴⁷

Das leituras realizadas verificamos que utilização da análise de citação como indicador de visibilidade, impacto e influência científica coloca muitas questões quer metodológicas quer éticas particularmente se se considerar como verdadeiro que o autor pode inflacionar, de forma artificial, o número de citações relativas a trabalhos seus.

Igualmente, alguns autores consideram que a utilização de indicadores bibliométricos como a análise de citações tem sido feita de forma a confundir a avaliação do impacto com a avaliação da qualidade por se entender que o interesse demonstrado por uma publicação, e traduzido pelas citações que lhe são feitas, é revelador da sua qualidade. Mas porque impacto e qualidade não são sinónimos é necessário distinguir ente qualidade, importância e impacto. A qualidade é uma característica intrínseca ao trabalho enquanto a importância e o impacto resultam de apreciações externas. As citações não constituem um indicador sobre a qualidade de uma publicação científica mas reflectem, contudo, o impacto que tem na comunidade científica.

Muitos autores incluem citações a trabalho seu em documentos posteriores. Existem várias explicações para que o façam, as quais foram já apresentadas. Este comportamento não tem estado isento a críticas por muitos considerarem que constitui uma forma artificial do os autores reforçarem o seu posicionamento num determinado

⁴⁷ Entre essas limitações podemos referir as pressões sociais e políticas existentes no seio da comunidade científica, a subjectividade dos avaliadores, a dificuldade em aceder à informação integral, o efeito de “halo”, o efeito de relações interpessoais positivas ou negativas, o custo elevado e a aplicação a “populações” pequenas.

domínio de investigação. Por esta razão, muitos consideram que as auto citações deveriam ser eliminadas quando são contabilizadas as citações que um determinado documento recebeu. Da leitura que fizemos do conjunto de artigos incluídos neste trabalho concluímos que os vários autores consideram que a taxa de auto citação face ao total de citações se enquadra em valores considerados aceitáveis (sobretudo a níveis de análise macro) sendo sugerido que, para esbater efeitos eventualmente perversos da auto citação, a análise de citação (incluindo auto citação) seja feita complementando-a com dados relativos ao factor de impacto das revistas (fornecidos pela base de dados ISI). De igual modo, o número de auto citações é mais elevado nos três anos seguintes à publicação do documento decaindo à medida que o tempo passa.

A forma como se desenvolve a investigação científica bem como as modalidades de comunicação dos resultados alcançados devem ser tomadas em consideração quando se utilizam os indicadores bibliométricos para avaliar a visibilidade, o impacto e a influência científica. A actividade de investigação tem, nos nossos dias, um pendor crescentemente colaborativo, seja à dimensão regional, nacional ou transnacional. A comunicação dos resultados continua a ser feita em publicações científicas impressas mas, também de forma crescente, é feita *online*, por exemplo em revistas em acesso livre, muitas delas com *peer review*.

Desta forma, são de prever alterações nas metodologias bibliográficas e bibliométricas à medida que cresce a disponibilização de revistas científicas e de trabalhos académicos na Internet. Desta forma “The presence of every scholarly work ever written linked to every work it cites or is cited by in a universal Web based bibliographic and citation database would solve many of the problems plaguing the construction of output and citation measures in today’s non electronic environment.”⁴⁸

Igualmente existe a expectativa de a fraca presença, em bases de dados científicas internacionais, de cientistas provenientes de países cuja língua materna não é o inglês, ser solucionada através da tradução automática, aumentando assim, os níveis de visibilidade e de citação das suas publicações.

É já uma realidade a existência de uma área de avaliação da comunicação electrónica designada Cibermetria (*Cybermetrics*) ou Webometria (*Webometrics*) discutindo-se a possibilidade e adequação da utilização dos métodos tradicionais na construção e produção de informação sobre citação usando o formato electrónico. Neste sentido, existem duas questões a que é necessário responder e que, por essa razão, constituem linhas de investigação futura: definir o que constitui uma publicação de qualidade na Internet e saber como calcular factores de impacto na Internet e a sua validade em termos de medição do impacto de uma determinada publicação.

Um aspecto que permanecerá inalterável é o elemento humano motivo pelo qual “Our poor understanding of the psychological and social processes, will not be automatically improved by wider access to output and impact data. Notwithstanding we will be able to provide peers with more comprehensive and more reliable bibliometric data to guide and support their decisions and to better defend these before different non-scientific sectors of the community, such as science managers, politicians⁴⁹ and the general public.”

A Bibliometria permite-nos medir para avaliar. E avaliamos com vários objectivos, como vimos ao longo deste trabalho. A utilização de indicadores de citação ganhou

⁴⁸ Russell, J. M. e Rousseau, R. (200?) – *Bibliometrics and institutional evaluation*. [Em linha]. p. 18.

Disponível em

<http://www.vub.ac.be/BIBLIO/itp/lecturers/ronald_rousseau/ronal_rousse_stim1_bibliometrics_russell.pdf>

⁴⁹ Op. cit., 19

uma dimensão nova quando os indicadores bibliométricos, para além de serem usados para avaliar o desempenho da investigação institucional e nacional, passaram a integrar a fórmula de financiamento da investigação científica.

Devemos ter presente que a avaliação da investigação não constitui um fim em si mesmo. Os constrangimentos de ordem financeira constituem, em parte, o fundamento para a necessidade da avaliação dado ser necessário distribuir os recursos disponíveis de forma equitativa.

Desta forma, para além da aplicação das metodologias e técnicas da Bibliometria na avaliação da investigação de forma a acompanhar as rápidas alterações que ocorrem nos padrões e nas práticas da comunicação científica importa igualmente “...constantly improve the theoretical foundation for the construction of output and impact indicators as an adjunct for peer review.”⁵⁰

⁵⁰ Russell, J. M. e Rousseau, R. (200?) – *Bibliometrics and institutional evaluation*. [Em linha]. p. 18.

Disponível em

<http://www.vub.ac.be/BIBLIO/itp/lecturers/ronald_rousseau/ronal_rousse_stim1_bibliometrics_russell.pdf>

Bibliografia

*Aksnes, D. W. (2003) – A macro study of self-citation. *Scientometrics*. Vol. 56 (2), 235-246. ISSN 0138-9130.

(citado 26 vezes sendo 1 delas pelo autor; contém 17 referências bibliográficas)

Cronin, B. (2001) – Hyperauthorship: a post modern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*. Vol. 52 (7), 558-569. ISSN 1532-2882.

(citado 17 vezes sendo 3 delas pelo autor – 1 sózinho e 2 em co-autoria; contém 98 referências bibliográficas)

Garfield, E. (1998) – Random thoughts on citationology: its theory and practice. *Scientometrics*. Vol 43 (1). ISSN 0138-9130.

(citado 16 vezes sendo uma delas pelo próprio; contém 30 referências bibliográficas)

*Glanzel, W., Thijs, B. e Schlemmer, B. (2004) – A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*. Vol. 59 (1), 63-77. ISSN 0138-9130.

(citado 18 vezes sendo 8 delas pelo autor em co-autoria; contém 7 referências bibliográficas)

*Glanzel, W. e Thijs, B. (2004) – Does co-authorship inflate the share of self-citations. *Scientometrics*. Vol. 61 (3), 395-404. ISSN 0138-9130.

(citado 2 vezes sendo duas pelo próprio em co-autoria; contém 7 referências bibliográficas)

*Glanzel, W., Debackere, K. e Schubert, B. T. (2006) – A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics*. Vol. 67 (2), 263. ISSN 0138-9130.

(citado 1 vez pelo autor em co-autoria; contém 25 referências bibliográficas)

*Lawani, S. (1981) – On the heterogeneity and classification of author self-citations. *Journal of the American Society for Information Science*. Vol. 33 (5), 281-284. ISSN 0002-8231.

(citado 17 vezes sendo 1 delas pelo próprio; contém 11 referências bibliográficas)

*MacRoberts, M. H. and MacRoberts, B. R. (1988) – Problems of citation analysis: a critical review. *Journal of the American Society for Information Science*. Vol. 40 (5), 342-349. ISSN 0002-8231.

(citado 140 vezes sendo 1 delas pelo próprio e 3 em co-autoria; contém 66 referências bibliográficas)

*Pichappan, P. and Sarasvady, S. (2002) – The other side of the coin: the intricacies of author self-citations. *Scientometrics*. Vol. 54 (2), 285-290. ISSN 0138-9130.

(citado 6 vezes sempre por outros; contém 20 referências bibliográficas)

Russell, J. M. e Rousseau, R. (200?) – *Bibliometrics and institutional evaluation*. [Em linha]. p. 18. Disponível em <http://www.vub.ac.be/BIBLIO/itp/lecturers/ronald_rousseau/ronal_rousse_stim1_bibliometrics_russell.pdf>

Warner, J. (2000) – A critical review of the application of citation studies to the Research Assessment Exercises. *Journal of Information Science*. Vol 26 (6), 453-460. ISSN 0165-5515.

(citado 14 vezes sendo 1 delas pelo próprio; contém 23 referências bibliográficas)

White, H. D. (2001) – Authors as citers over time. *Journal of the American Society for Information Science and Technology*. Vol 52 (2), 87-108. ISSN 1532-2882.

(citado 24 vezes das quais 2 pelo próprio e 2 em co-autoria; contém 85 referências bibliográficas)