

# RESEARCH METHODOLOGY

QUANTIFYING  
RESEARCH  
OUTPUT



1

A researcher (say X) produced 25 publications on a problem and had a total of 10 citations. Another one (say Y) produced 5 publications on the same problem and the total number of citations for these 5 papers was 50. In your opinion, which of these researchers had made more significant contribution to the field?



# HOW TO QUANTIFY RESEARCH OUTPUT?

- Total number of papers
- Total number of citations
- Average number of citations per paper
- Average number of citations per author
- Average number of citations per author per year
- Average number of papers per author
- Average number of authors per paper
- Some thing else?



How would I know if others are citing my work?

# PROFILE – GOOGLE SCHOLAR



**Imran Siddiqi**

Bahria University Islamabad  
Image Analysis and Pattern Recognition  
Verified email at bahria.edu.pk - [Homepage](#)

[Follow](#)

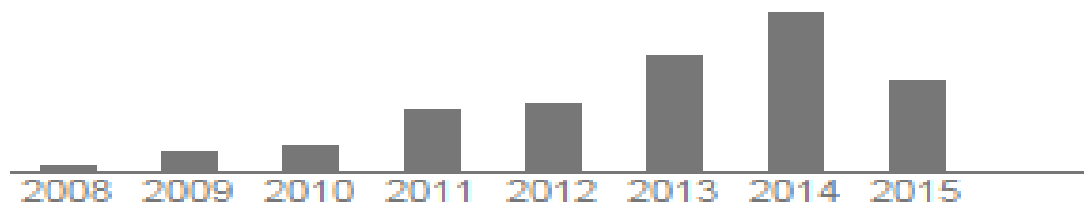
Title	1–20	Cited by	Year
<a href="#">Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features</a>	I Siddiqi, N Vincent Pattern Recognition 43 (11), 3853-3865	58	2010
<a href="#">Comparison of Niblack inspired Binarization methods for ancient documents</a>	K Khurshid, I Siddiqi, C Faure, N Vincent 16th Document Recognition and Retrieval Conference DRR	55	2009
<a href="#">A set of chain code based features for writer recognition</a>	I Siddiqi, N Vincent Document Analysis and Recognition, 2009. ICDAR'09. 10th International ...	53	2009
<a href="#">Writer identification in handwritten documents</a>	IA Siddiqi, N Vincent Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International ...	37	2007
<a href="#">Combining global and local features for writer identification</a>	I Siddiqi, N Vincent Proceedings of the 11. Int. Conference on Frontiers in Handwriting ...	24	2008
<a href="#">Text-independent writer recognition using multi-script handwritten texts</a>	C Djeddi, I Siddiqi, L Souici-Meslati, A Ennaji	21	2013

# PROFILE - GOOGLE SCHOLAR

Google Scholar

[Get my own profile](#)

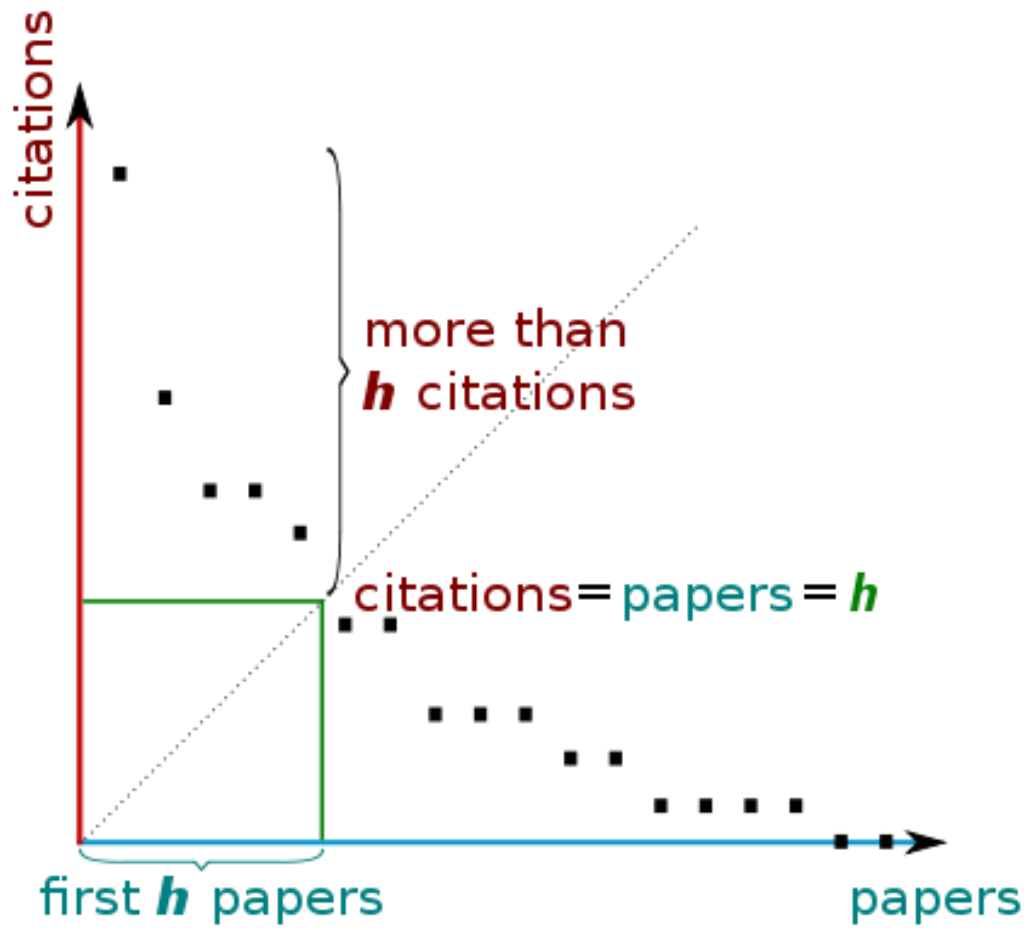
Citation indices	All	Since 2010
Citations	338	315
h-index	9	9
i10-index	9	9



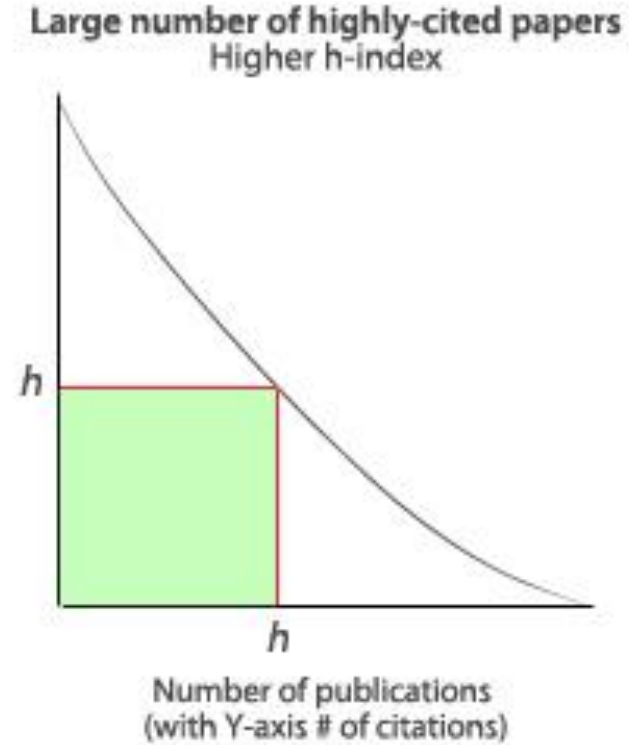
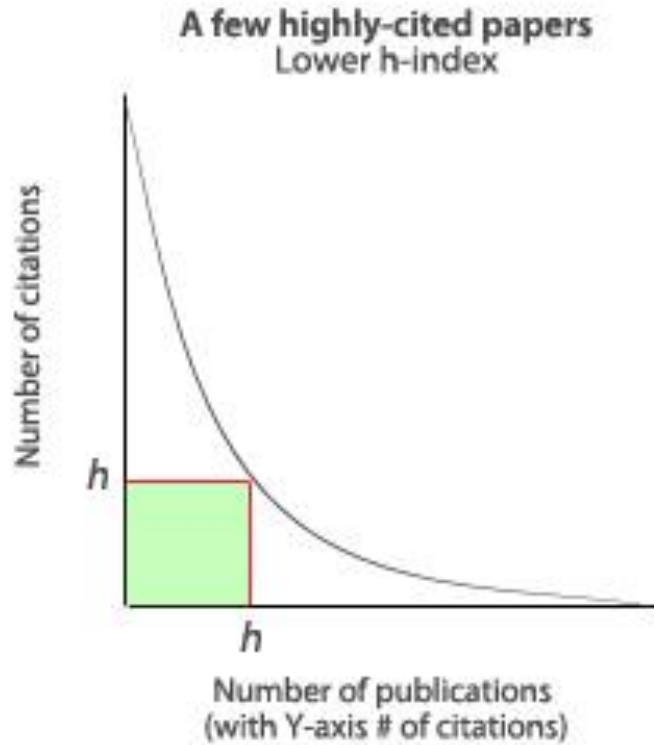
# H-INDEX

- The h-index, originally described in 2005 by Jorge Hirsch
- The ***h-index*** is an index that attempts to measure both the productivity and impact of the published work of a scholar
- *The h-index of a scholar is  $k$  if the scholar has  $k$  publications with at least  $k$  citations each*
- *h-index* reflects both the number of publications and the number of citations per publication
- The index is designed to improve upon simpler measures such as the total number of citations or publications.

# H-INDEX



# COMPARING H-INDICES



# H-INDEX

- What is good?
- Varied opinions



After 20 years a "*successful scientist*" will have an *h*-index of **20**, an "*outstanding scientist*" an *h*-index of **40**, and a "*truly unique*" individual an *h*-index of **60** [Hirsch]

## H-INDEX: PROS AND CONS

- It is not severely twisted by a single well-cited, important paper (unlike total number of citations would be)
- It is not increased by a large number of poorly cited papers (unlike total number of papers would be)
- It limits authors by the total number of publications, so shorter careers are at a disadvantage
- It has relatively low resolution in that many scientists end up in the same range since it gets increasingly difficult to increase the h-index the higher it gets (an h-index of 100 corresponds to a minimum of 10,000 citations)

# TOP COMPUTER SCIENTISTS – IN TERMS OF H-INDEX

<http://www.cs.ucla.edu/~palsberg/h-number.html>

- 151 [Herbert A. Simon](#) (CMU), [Nobel Laureate](#), [Turing Award](#), [ACM Fellow](#)
- 139 [Anil K. Jain](#) (Michigan State U), [ACM Fellow](#), [IEEE Fellow](#), [IEEE Technical Achievement Award](#)
- 120 [Terrence Sejnowski](#) (UCSD), [IEEE Fellow](#), [Member of the National Academy of Engineering](#), [Member of the](#)
- 119 [Scott Shenker](#) (Berkeley), [ACM Fellow](#), [IEEE Fellow](#), [Member of the National Academy of Engineering](#)
- 118 [Hector Garcia-Molina](#) (Stanford), [ACM Fellow](#), [Member of the National Academy of Engineering](#)
- 118 [Takeo Kanade](#) (CMU), [ACM Fellow](#), [IEEE Fellow](#), [Member of the National Academy of Engineering](#)
- 114 [Jiawei Han](#) (UIUC), [ACM Fellow](#), [IEEE Technical Achievement Award](#)
- 111 [Tomaso Poggio](#) (MIT)
- 110 [David Haussler](#) (UC Santa Cruz)
- 110 [Sebastian Thrun](#) (Stanford), [Member of the National Academy of Engineering](#)
- 109 [Ian Foster](#) (Argonne National Laboratory & U Chicago), [ACM Fellow](#), [AAAS Fellow](#), [BCS Fellow](#)
- 108 [Michael I. Jordan](#) (Berkeley), [ACM Fellow](#), [IEEE Fellow](#), [AAAS Fellow](#), [AAAI Fellow](#), [Member of the National Academy of Arts & Sciences](#), [SIAM Fellow](#)
- 106 [Andrew Zisserman](#) (University of Oxford), [Fellow of the Royal Society](#)
- 105 [Wil van der Aalst](#) (TU Eindhoven), [Member of Academy of Europe \(Academia Europaea\)](#), [Member of the R](#)
- 105 [Philip S. Yu](#) (UIC), [ACM Fellow](#), [IEEE Fellow](#), [IEEE Technical Achievement Award](#)
- 104 [Deborah Estrin](#) (UCLA and Cornell NYC Tech), [ACM Fellow](#), [IEEE Fellow](#), [Member of the National Acad](#)
- 104 [Christos H. Papadimitriou](#) (Berkeley), [ACM Fellow](#), [Member of the National Academy of Engineering](#), [Mem](#)
- 101 [Alex Pentland](#) (MIT)
- 101 [Steven Salzberg](#) (Johns Hopkins U), [AAAS Fellow](#)
- 100 [Georgios Giannakis](#) (U Minnesota), [IEEE Fellow](#)
- 100 [Nick Jennings](#) (U Southampton), [IEEE Fellow](#), [Fellow of the Royal Academy of Engineering](#), [Member of Acc](#)
- 97 [Jack Dongarra](#) (U Tennessee), [ACM Fellow](#), [IEEE Fellow](#), [Member of the National Academy of Engineering](#),
- 95 [Didier Dubois](#) (CNRS Toulouse)

# G- INDEX

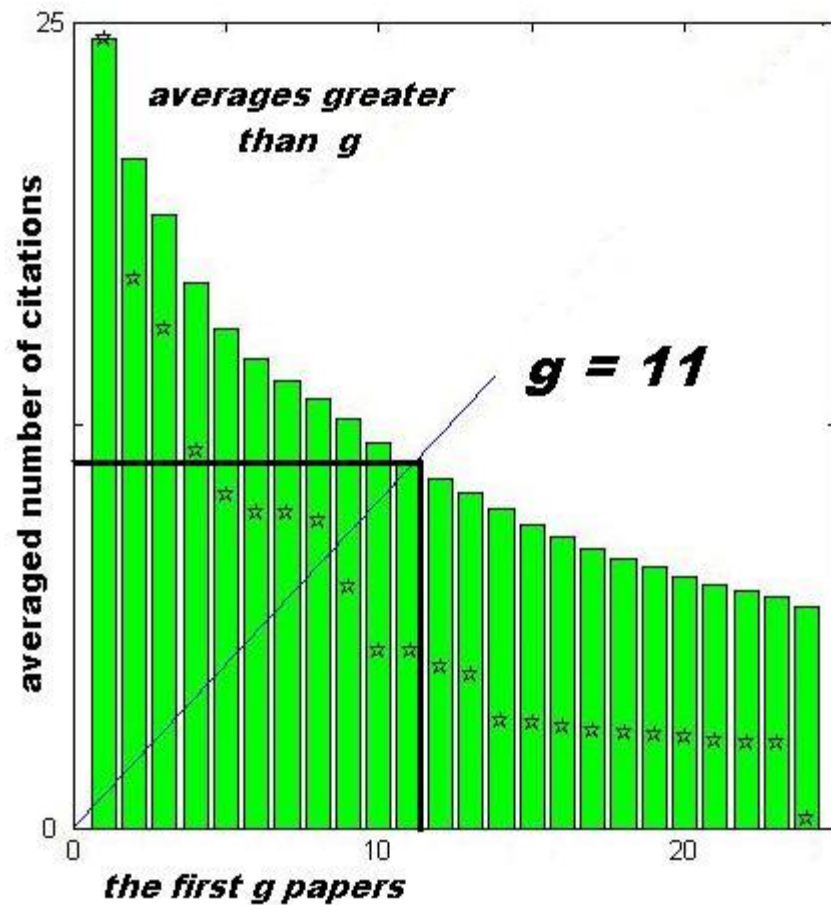
- The g-index was proposed by Leo Egghe in 2006
- [Given a set of articles] ranked in decreasing order of the number of citations that they received, the g-index is the largest number such that the top g articles received (together) at least  $g^2$  citations.
- It aims to improve on the h-index by giving more weight to highly-cited articles.

# g-index for Professor X

The top g articles received (altogether) at least g squared citations.

Document no. (g)	Citation count	Square of g	Total no. of citations
Document 1	50 cites	1	50
Document 2	18 cites	4	50+18 = 68
Document 3	11 cites	9	68+11 = 79
Document 4	7 cites	16	79+7 = 86
Document 5	4 cites	25	86+4 = 90
Document 6	3 cites	36	90+3 = 93
Document 7	0 cites	49	93
Document 8	0 cites	64	93
Document 9	0 cites	81	93
Document 10	0 cites	100	93

# G-INDEX



$$g^2 \leq \sum_{i \leq g} C_i$$

# G-INDEX

- In order to have a g-index of  $n$  an author that produces  $n$  articles should have, on average,  $n$  citations for each of them.
- Unlike the h-index, the g-index depends on the full citation count of very highly cited papers.
- In all cases  $g$  is at least  $h$ , and is in most cases higher

# I10 -INDEX

- *Introduced by Google in 2011*
- *i10-index is the number of publications with at least 10 citations*
- *Not very popular*

# REFERENCES

- Google Scholar
- The h Index for Computer Science, by Jens Palsberg
- H-Index: What It Is and How to Find Yours by Alan Marnett