What if honest error is reason for retraction ?

Initiators and reasons for retraction

Thed van Leeuwen & Jochem Zuijderwijck

Centre for Science and Technology Studies (CWTS), Leiden University

Retractions (CC16),
Tuesday, June 4th,
Room CPD-3.04, 3/F, Run Run Shaw Tower, Centennial Campus, HKU

WCRI 2019 Conference, Hong Kong, China
June 2-5th, 2019

H2020 project Printeger: On retractions

- We assumed to find many reasons for retractions:
  - Plagiarism;
  - Data falsification & manipulation;
  - Data fabrication;
  - Sloppiness with respect to ethical issues;
  - Twice or duplicate publishing of the same results;
  - Common errors by authors or publishers;
  - etc. …

- Due to a number of highly publicized cases, scientific fraud came under scrutiny of academic authorities, funding agencies, the media, politicians, and the public at large.

CWTS

Retractions in journals covered by WoS

- Several studies on retractions have been published, mostly focused on clinical sciences (PUBMED) or samples of retracted papers in WoS journals (see e.g. Fang et al. (2012), Chen et al. (2013))

- We reported in the project the first results of an ongoing study of **_all_** retracted papers published in journals processed for the WoS

  ➔ in December 2016, we downloaded **3.729** papers

- Be beware, we talk about an extremely small fraction:
  ➔ only 0,008% of WoS articles

CWTS

Data collection

- **Retracted paper in WoS** – in the title of the paper the bibliographic information of the retraction notice is given

'Transplanted iNSCs migrate through SDF-1/CXCR4 signaling to promote neural recovery in a rat model of spinal cord injury (Retracted article. See vol. 25, pg. 806, 2014)'

- The **Retraction notification** contains information on the party responsible for the retraction, and its <u>motivation</u>.

- Pairs of retracted papers & retraction notices (pdf file)
  - Manually collected & analyzed the pdf files – tedious and time consuming work

Motivation for the retraction

	Motivation	Nr of retractions	Share %
1	unknown (no motivation given)	527	14%
2	errors	391	10%
3	duplicate publishing	386	10%
4	plagiarism	295	8%
5	data falsification	258	7%
6	data irreproducible	169	5%
7	unknown (no full text available)	154	4%
8	self plagiarism	142	4%
9	peer review issues	131	4%
10	data fabrication	122	3%
11	data inaccuracies-inconsistencies-irregularities	110	3%
12	ethical issues	110	3%
13	authorship issues	108	3%
14	unknown (not found on journal website)	108	3%
15	unknown (no license agreement with publisher)	98	3%

Motivation for the retraction

	Motivation	Nr of retractions	Share %
1	**unknown (no motivation given)**	**527**	**14%**
2	errors	391	10%
3	duplicate publishing	386	10%
4	plagiarism	295	8%
5	data falsification	258	7%
6	data irreproducible	169	5%
7	**unknown (no full text available)**	**154**	**4%**
8	self plagiarism	142	4%
9	peer review issues	131	4%
10	data fabrication	122	3%
11	data inaccuracies-inconsistencies-irregularities	110	3%
12	ethical issues	110	3%
13	authorship issues	108	3%
14	**unknown (not found on journal website)**	**108**	**3%**
15	**unknown (no license agreement with publisher)**	**98**	**3%**

Motivation for the retraction:
The 'unknown' factor

	Motivation	Number of retractions	Share %
1	unknown (no motivation given)	527	14%
7	unknown (no full text available)	154	4%
14	unknown (not found on journal website)	108	3%
15	unknown (no license agreement with publisher)	98	3%

- **Unknown reason for Retraction**– some 25% of all retractions, makes one wonder about taboos in academic publishing!

CWTS

Motivation for the retraction: the expected reasons

Motivation	Number of retractions	Share %
2 errors	391	10%
3 duplicate publishing	386	10%
4 plagiarism	295	8%
5 data falsification	258	7%
6 data irreproducible	169	5%
8 self plagiarism	142	4%
9 peer review issues	131	4%
10 data fabrication	122	3%
11 data inaccuracies-inconsistencies-irregularities	110	3%
12 ethical issues	110	3%
13 authorship issues	108	3%

CWTS

On honest errors (HE)

- The working hypothesis is: ***honest error should in principle not harm an academic career !***

  - ➢ … as authors seem to be the main initiating party in retracting these erroneous publications, while …

  - ➢ … in FFP-related cases, editors and/or publishers take the initiative !

- Work with the 391 papers, retracted due to honest error (HE)

- Collect all authors connected to the retracted papers, both connected to FFP, QRP, as well as to HE

- From the CWTS author database, calculate length of careers, and put that into the perspective of the retraction(s)

CWTS

Data and indicators used

- For every author and her/his oeuvre we know:
  - *Year of first publication*
  - *Year of last publication*
  - *Year retraction took place*

  Length of academic career

- Exclusion of authors with 'tricky' names
  - Common Western names (Smith W, Brown D, Moller A, Visser M, etc.)
  - Korean and Chinese names (Park, Zhang, Wu, etc.)

- Focus on academic age and career length
  - ***Early career***: whenever retractions <u>occur</u> in first 5 years of an academic career
  - ***Midrange***: whenever retractions <u>occur</u> in careers up to 15 years of length
  - ***Long***: whenever retractions <u>occur</u> after 15 years of length
  - ***Year-1***: whenever a retraction <u>occurs</u> in the first year of academic publishing

- Focus on retractions **<u>occurring</u>** in an early career stage
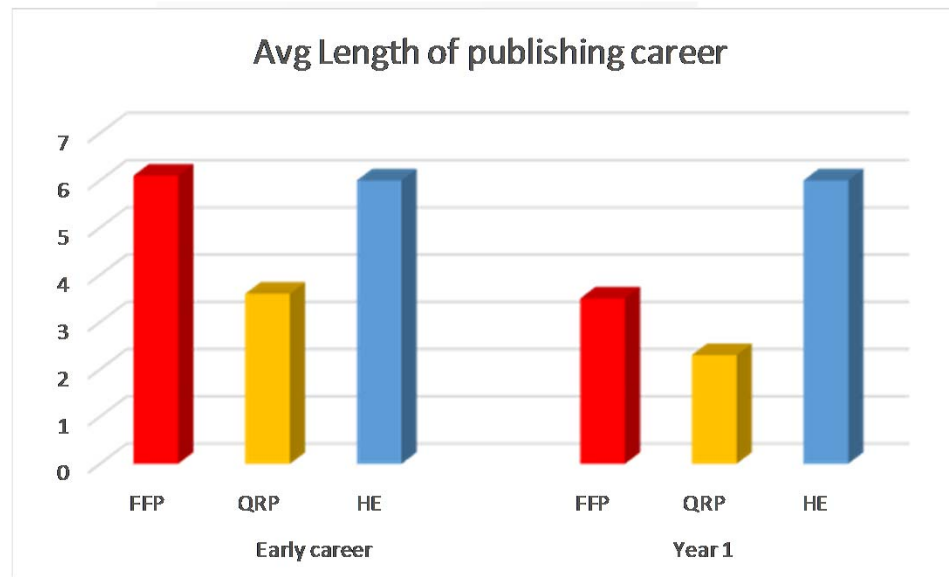 - Focus on ***Early career*** and ***Year-1*** as characteristics

1- Results: how are authors distributed ?

|	Early Career	Year-1	Midrange	Long	Total authors	Total publications	Mean AU/PU
FFP	2117	*1073*	1470	1744	5331	1051	5,1
QRP	2808	*1598*	1584	1427	5819	1035	5,6
HE	948	*457*	818	924	2690	391	6,9
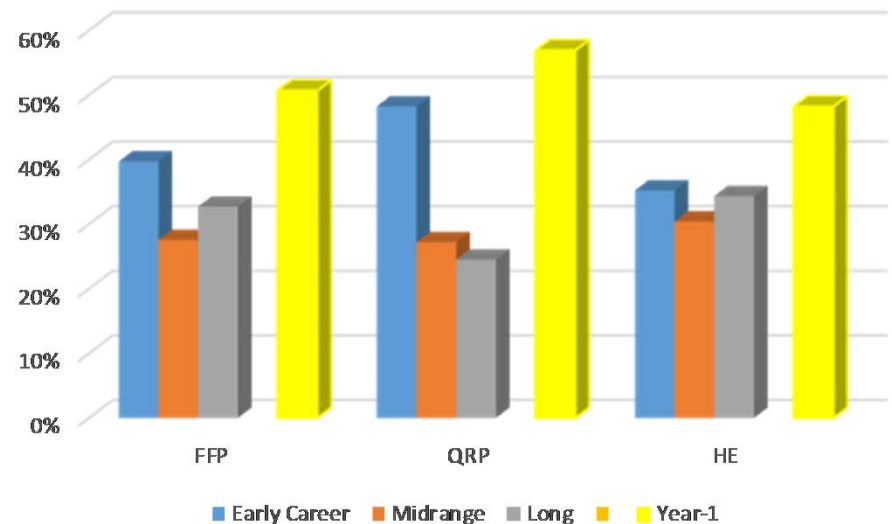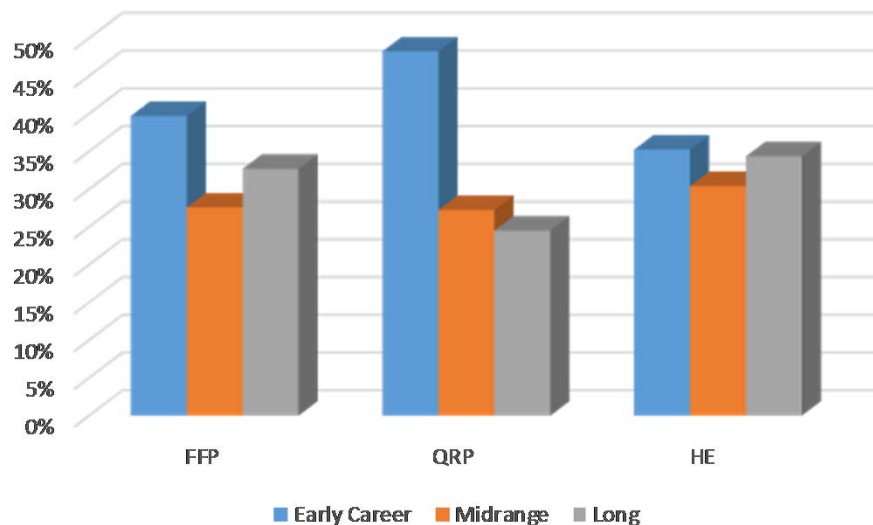
2 - Results: average effect on career length

		Avg nr Years before retraction	Avg nr Years publishing after retraction	Avg Length of publishing career
Early career	FFP	2,1	3,9	6,1
	QRP	2,0	3,2	3,6
	HE	2,2	3,7	6,0
Year 1	FFP	1,0	2,5	3,5
	QRP	1,0	1,7	2,3
	HE	1,0	1,8	6,0

Avg Length of publishing career

3 - Results: Devastating effect on early career length when retraction occurs in Year-1

		year-1	yrs 2-5	yrs 6-10	yrs > 10
Early career	FFP	629	841	489	158
	QRP	967	1166	528	147
	HE	364	324	161	99
Year 1	FFP	857	564	131	46
	QRP	479	355	183	56
	HE	266	114	64	13
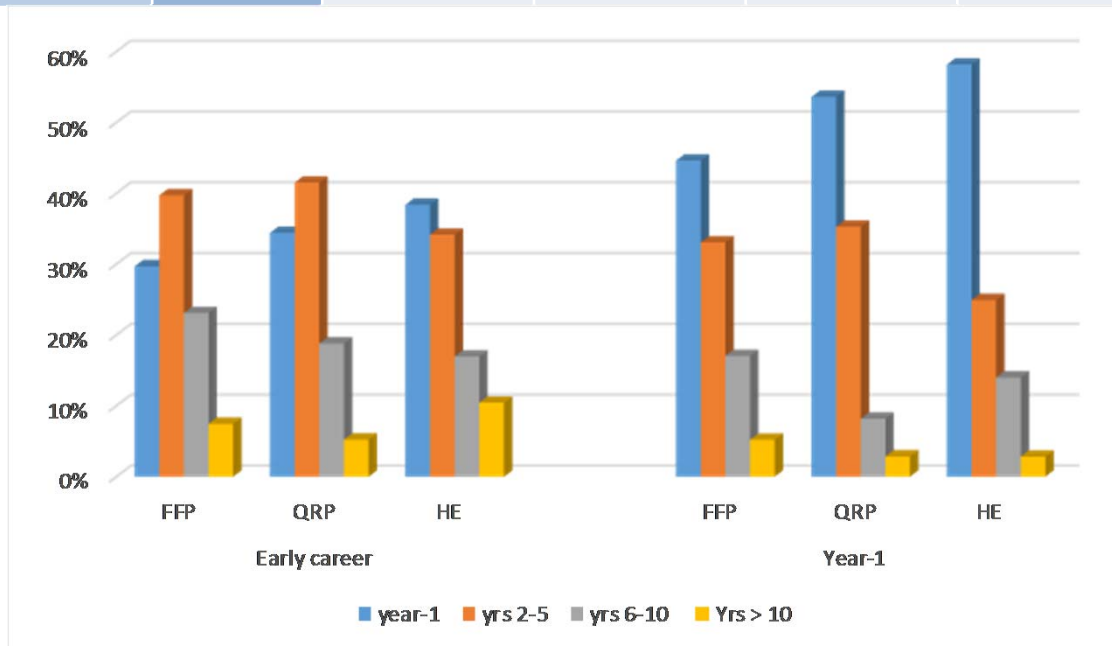
Conclusions

- With **FFP** and **QRP**, consequence of early occurrence of retraction seems to shorten research careers.

- In particular when this occurs in the first year of activity, effects seem to be devastating (length of career decreases by 40% !)

- With **HE**, occurrence of retraction does <u>on average </u>not seem to influence the length of a research career

- Moment of occurrence is very relevant, as occurrence in Year-1 means immediate shortening of research career.

- Particularly in **HE**, as some 60% of all authors seem to drop out of academic publishing !

CWTS

Next steps and discussion

- These findings need further exploration
    - Geographical orientation
    - Disciplinary orientation
    - Position of Early career drop outs on author list

- **Also**: conduct interviews to compare these outcomes with !

- At the minimum, the study shows the strength of bibliometrics to quantitatively support and inform these type of analyses on retraction and career effects.

CWTS

Thank you for your attention !

For questions, ask me now,
or mail us...

leeuwen@cwts.nl
j.b.zuijderwijk@cwts.leidenuniv.nl.

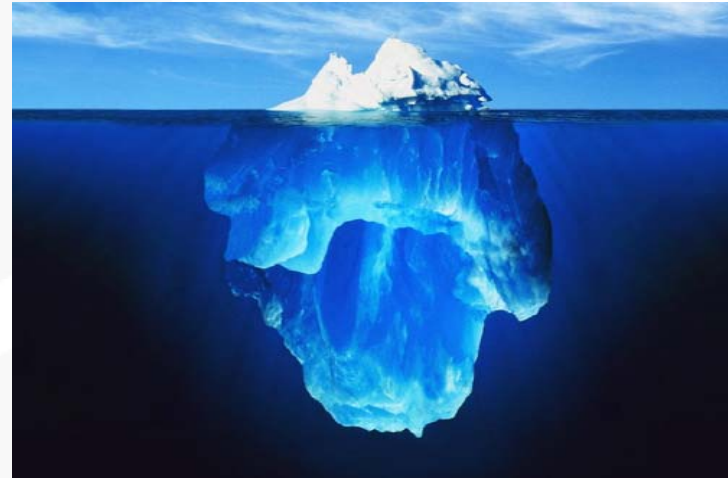# The Iceberg metaphor

Perspectives on scientific misconduct

- Misconduct is often discussed in terms of metaphors

- Metaphor of the *"Bad apple in the barrel"*
    - Misconduct is individual researcher driven
    - Incidental behavior

- Metaphor of the *"Iceberg-model"*
    - Misconduct is a systemic characteristic of science
    - Occurrences are weaving errors

Consequences of "Iceberg metaphor"

- The Iceberg model says that about 10% of the Iceberg is above the water surface, while some 90% is below the surface.

- In this study, we worked with roughly 4.000 publications, being retracted *(the 10% of our iceberg)*, which means that we have no clue of about another possibly 36.000 affected publications in WoS *(the other 90% of our iceberg)*.

How about these mind boggling numbers?

- From the systemic interpretation, we can consider WoS as our iceberg.

- So the 'iceberg' counts over 50 million publications.

- Below the water surface, there are according to the model's distribution, some 45 million publications

- Above the water surface, there should be 5 million publications.

- According to this reasoning, the number of retracted publications should be 5 million …

- … and we found only some 4000 publications retracted in WoS ….

- ?????