

Terje Tüür-Fröhlich

# The Non-trivial Effects of Trivial Errors in Scientific Communication and Evaluation

**vwh**

Verlag Werner Hülsbusch  
Fachverlag für Medientechnik und -wirtschaft

T. Tüür-Fröhlich: The Non-trivial Effects of Trivial Errors in Scientific Communication and Evaluation

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://d-nb.de> abrufbar.

© Verlag Werner Hülsbusch, Glückstadt, 2016

**vwh** Verlag Werner Hülsbusch  
Fachverlag für Medientechnik und -wirtschaft

[www.vwh-verlag.de](http://www.vwh-verlag.de)

Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Glückstadt.  
Eine weitere Verwertung im Sinne des Urheberrechtsgesetzes ist nur mit Zustimmung der Autorin möglich.

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

Satz: Werner Hülsbusch  
Druck und Bindung: SOWA Sp. z o. o., Piaseczno

Printed in Poland

Zugleich: Diss., Sozial- und Wirtschaftswissenschaftliche Fakultät der Johannes Kepler Universität Linz, 2014

– Als Manuskript gedruckt –

ISSN: 0938-8710

ISBN: 978-3-86488-104-6

# Contents

	<b>Preface</b> ( <i>Volker Gadenne</i> )	<b>5</b>
	<b>Acknowledgements</b>	<b>7</b>
	<b>Contents</b>	<b>9</b>
	<b>PART ONE: TRIVIAL ERRORS &amp; COMMERCIAL SCIENCE EVALUATION</b>	
<b>1</b>	<b>Introduction</b>	<b>15</b>
1.1	Trivial Errors in Doctoral Students' Everyday Life	15
1.2	Basic Scientometrics: Technical Terms Used	17
1.3	Scientific Ethos, Matthew & Matilda Effects: Selected Concepts of Science Studies	20
1.3.1	Scientific Ethos (Robert K. Merton)	20
1.3.2	Matthew Effects (Robert K. Merton)	22
1.3.3	Matilda Effects (Margaret Rossiter)	22
1.4	Science Policy Issues and Academic Springs: Protests Against the Quantitative Evaluation of Scientific Achievements	23
1.4.1	DORA: "Putting Science into the Assessment of Research"	24
1.4.2	Protests Against Rankings and Monopolistic Power Positions	26
1.4.3	Alternative Forms of Critical Scientific / Scholarly Communication	27
<b>2</b>	<b>Aims and Scope of my Study</b>	<b>29</b>
2.1	Sciences As Error Making Activities (Popper)	29
2.2	Research Theses: The High Relevance of Trivial Errors in Scientific Evaluation	30
2.3	Methods	31
2.4	Limitations of Errors Research Literature	31
2.4.1	Error Managment Literature	32
2.4.2	"Typos" and "accuracy of references"	32
2.4.3	Financial Analyses Literature	34
2.4.4	Name Disambiguation Literature	34

<b>3</b>	<b>Global Public Science Evaluation Is Based on Privately Owned Data</b>	<b>37</b>
3.1	Thomson Reuters' Influence on University Rankings	37
3.1.1	Times Higher Education World University Ranking	38
3.1.2	Other International Rankings influenced by Thomson Reuters' Data	39
3.2	General Evaluation Effects	40
3.2.1	The Gratification of the Chosen	40
3.2.2	The Reactivity of Rankings	43
3.3	Specific Evaluation Effects: Trivial Errors in Thomson Reuters' Data and their Effects	44
<b>PART TWO:</b>		
<b>ERRORS IN THOMSON REUTERS' SSCI – QUALITATIVE AND QUANTITATIVE CASE STUDIES</b>		
<b>4</b>	<b>Errors in Thomson Reuters' SSCI (Social Sciences Citation Index) – Qualitative and Quantitative Case Studies</b>	<b>49</b>
4.1	Pierre BOURDIEU as Cited Author in the <i>Social Sciences Citation Index</i> (SSCI)	50
4.1.1	The Choice of Pierre Bourdieu as Subject of a Case Study	50
4.1.2	Methods: Data Elements and “Ping-Pong Method”, “Snow-Ball Sampling”	50
4.1.3	Results	52
4.1.4	“ <i>pierre b</i> ” and “ <i>pierr b</i> ”: The Continuous Confusion of Surnames and Given Names – Unknown in the Scientometric Community?	57
4.2	The Representation of Séverine SOFIO (2008, <i>Travail, genre et sociétés</i> ) in SSCI	57
4.2.1	Cited Author: “PIERRE B”	57
4.2.2	Modus Operandi of SSCI Record Analyses: Indexing Rules and the Hardships of Comparisons (ORIGINAL with RECORD)	58
4.2.3	Elaborating the Problems of Comparison SOFIO (2008) Original with SSCI Record	60
4.2.4	Results	61
4.3	Quantitative Case Study: The Representation of Barton BEEBE (2010, <i>Harvard Law Review</i> ) in SSCI	64
4.3.1	Cited Author: “PIERRI B”	64
4.3.2	Modus Operandi, the <i>Bluebook</i> Citation Style and Matching Problems	65
4.3.2.1	<i>The Bluebook</i> as Legal System of Citation in the USA	65
4.3.2.2	BEEBE 2010, ORIGINAL versus RECORD: Identifying Matching Pairs	67
4.3.2.3	Data Elements: “whatever can go wrong, will go wrong”	68

4.3.3	Synopsis Results: Losses, Severe Errors (Mutations/Mutilations/Zombies), Minor Errors, Correct References in the SSCI record	68
4.3.4	Results: <i>Missing Items</i> in the SSCI Record (56%)	69
4.3.4.1	Doublets and Triplets in the SSCI record	70
4.3.4.2	An UBO Turned Out to Be a Cuckoo's Egg	74
4.3.5	Results: Multiple Severe Errors (40%)	75
4.3.5.1	Variants of Phantom Authors by Genesis	75
4.3.5.2	Variants of Phantom References by Genesis (Error Types 'Lumped Together')	86
4.3.5.3	Zombie References: Anything Is Wrong But Something Has Been Indexed	91
4.3.6	Results: Items With Minor Errors (2,8%), Correct Items (1%)	100
4.3.6.1	"Minor Errors"	100
4.3.6.2	Correct Items	101
4.4	The Representation of Bruce CURTIS (2007, <i>Canadian Journal of Sociology/Cahiers canadiens de sociologie</i> ) in SSCI	103
4.4.1	Cited Author "Bourdieu", Document Type: "Book Review"	104
4.4.2	Results of Comparison ORIGINAL versus RECORD	104
4.5	Conclusions Case Studies	106
4.5.1	Case Study: The 'Indexing Fate' of Pierre Bourdieus's Name in SSCI/WoS	107
4.5.2	Quantitative Case Studies: The Fate of References of Severine Sofio (2008), Barton Beebe (2010) and of Reviewed Book/References of Bruce Curtis (2007)	108

### PART THREE: HISTORICAL ROOTS – DISCUSSION – CONCLUSIONS – RECOMMENDATIONS

<b>5</b>	<b>Fatal Data Errors in SSCI – Why? Intransparency of TR's Data Capture, Indexing and Error Correction/ Avoidance Methods</b>	<b>113</b>
5.1	TR's Data Input and Indexing Procedures: Vague and Conflicting Statements	113
5.1.1	Data Input: Automated, Semi-Automated, or by Human Indexers?	113
5.1.2	Two Classes of Journals by different Data Input Techniques?	115
5.1.3	"AGOS78 1 453" or Identifying Codes/Citation Tags as Cuckoo's Eggs?	116
5.1.4	The Accuracy of Text-Parsing	118
5.2	TR's Error Detection/Avoidance/Correction Procedures: Vague and Conflicting Hints/Information Fragments	119

5.2.1	Current TR Practices	119
5.2.2	Garfield's Statements on Quality Control / Error Management at ISI	120
5.2.4	Different Database Versions	123
<b>6</b>	<b>Fatal Data Errors in SSCI – Why? Strategies and Contingencies in the Genesis of Commercial Science Citation Indexing, an “Error-Making” (Popper) Activity</b>	<b>125</b>
6.1	The Strategy: Spreading Overoptimism, Downplaying Severe Problems	125
6.2	First Error Reports: “ <i>The standard of scientific publication in the US is rapidly deteriorating</i> ” (JBS Haldane) – “ <i>More A Comedy of Errors Than A Real Loss</i> ” (Lederberg)	127
6.3	Contingencies	128
6.3.2	The Emergence as Genetics Citation Index in the Punch Card Era	129
6.3.3	Vexed Problem of Getting Funded	130
6.4	The Inertia of Thomson Reuters' Commercial Citation Indexing	130
<b>7</b>	<b>Conclusions and Recommendations</b>	<b>133</b>
7.1	Conclusions: Research Theses and Accumulated Evidence	133
7.1.1	The High Relevance of Trivial Errors in Evaluation	133
7.1.2	Errors as Biases, Errors and Biases	133
7.1.3	“The Tomato Was Rotten From the Beginning”	136
7.2	Recommendations: “Putting Science into the Assessment of Research” (DORA)	137
	<b>Reference List</b>	<b>141</b>
	<b>Appendix</b>	<b>153</b>
	<b>Tables</b>	<b>161</b>
	Table of Tables	161
	Table of Lists	162