The Matthew effect and a relation with concept symbols and defaults

Dilruba Mahbuba\textsuperscript{1,2} and Ronald Rousseau\textsuperscript{2,3,4}

\textsuperscript{1}Library and Information Department, Northern University Bangladesh (NUB), Sher Tower, Holding \# 13, Road \# 17, Banani, Dhaka – 1213, Email: dilrubauw@gmail.com

\textsuperscript{2}University of Antwerp, IBW, Venusstraat 35, B-2000 Antwerp, Belgium

\textsuperscript{3}KHBO (Association K.U.Leuven), Industrial Sciences and Technology, Zeedijk 101, B-8400 Oostende, Belgium

\textsuperscript{4}K.U.Leuven, Dept. Mathematics, Celestijnenlaan 200B, B-3000 Leuven (Heverlee), Belgium

In this article a review of Merton’s article about the Matthew effect is presented. The authors discuss what they think Merton really meant and link it to the use of concept symbols and defaults. The Matthew effect is linked to the notions of concept symbols and defaults. The diffusion of Merton’s ideas in science over time is discussed. Several h- and R-indices related to “Matthew” publications are derived. Related effects, namely the Podunk effect, the Knudop effect and the Wehttam effect are briefly discussed. Also the Matilda effect is given some attention as is the under representation of women in science and factors explaining this phenomenon.

Keywords: Matthew effect; concept symbols, defaults; diffusion; women in science

Introduction

The Matthew effect according to which the rich get richer and the poor get poorer (a more precise statement follows) is an explaining principle in sociology and economics. From sociology and economics it has been introduced in the information sciences and linked to the cumulative advantage and success-breeds-success effects. Basically the Matthew effect is the occurrence of a positive feedback loop\textsuperscript{1}. This effect is the basic explanation of the ubiquity of power laws in informetrics and several other fields\textsuperscript{2}. In this article we present a review of Merton’s article about the Matthew effect. We discuss what we think Merton really meant and link it to the use of concept symbols and defaults. We further consider the diffusion of his ideas in science over time. Among those colleagues that have been inspired by Merton we give special attention to Manfred Bonitz. Several h- and R-indices related to “Matthew” publications are derived. We briefly discuss derived and related effects, namely the Podunk effect, the Knudop effect and the Wehttam effect, proposed by Gaston\textsuperscript{3}. Next we move on to the Matilda effect and briefly discuss the underrepresentation of women in science and factors explaining this. This article is a revised version of a contribution to the Collnet Meeting and WIS conference held in Istanbul (September 20-23, 2011) at Bilgi University\textsuperscript{4}.

The Matthew Effect

By his contribution in the journal \textit{Science} in 1968 entitled “The Matthew Effect in science” and subtitled “The reward and communication systems of science are considered”, Robert K. Merton made the term “Matthew Effect” well-known to a broad audience\textsuperscript{5}. In this section we present a synthesis of Merton’s famous article. From now on it will be denoted as MERT68. Simply put, the term “Matthew Effect” as used in sociology and the information sciences refers to the habit of giving credit to already famous people and minimizing or withholding recognition from scientists who have not (yet) made their mark. In the stately language of the bible, Merton wrote it as:

\begin{quote}
For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.
\end{quote}

Yet, as can be derived from the title of MERT68 the term “Matthew Effect” has already been used before,
namely in economics by James S. Duesenberry. Moreover, use of this term to describe this particular pattern of accumulation of prestige was suggested to Merton by Marshall Childs. So the fact that we think that the term “Matthew Effect” was invented by Merton is itself the result of a Matthew Effect.

In MERT68, Merton describes the Matthew Effect as a fact placed in the context of the reward system of science. As such he describes the Matthew Effect as a complex psychosocial process affecting the reward system as well as the communication system of science. According to Merton, it occurs mainly in cases of collaboration and in cases of multiple discoveries: in both cases the famous person receives all (or most of) the credit for the collaborative work or the multiple discovery (the term ‘multiple discovery’ is used here in the sense of: one discovery but done by two or more persons at about the same time). Famous persons probably also enjoy higher acceptance rates for their articles, yet this does not shield them from the ‘danger’ of receiving no citations.

According to Merton, the Matthew Effect is a form of optimization in the scientific communication system. Indeed, a scientific contribution or hypothesis has a higher visibility within the scientific community when it comes from a famous person, than when it comes from a (younger) person who is as yet unknown. We believe that instead of younger person it can also be ‘woman’, or ‘scientist from a developing country’ (see further). In this way the Matthew Effect may increase the visibility of interesting new ideas and hypotheses, and hence act to smooth the flow of scientific ideas. Another point in favour of the Matthew Effect is that it may act as a cue for reading. Indeed, scientists nowadays get buried under mountains of new publications, the so-called information overload. When famous persons are known for publishing only contributions that they think are worth the attention of others, this may act as a cue for choosing such articles for reading (or at least skimming through them). As Merton puts it so aptly: only work that is effectively perceived and utilized by other scientists ... matters. Making innovative ideas known is most likely (but of course not uniquely) done by scientists of high standing. They are the ones that are able to accelerate the development of science. Merton further discusses the function of redundancy. This is especially interesting for researchers of lesser esteem in the scientific system. Their ‘redundant’ contributions are necessary to make good ideas known, while this is less the case for important persons. Yet, and in particular, it is sometimes a ‘redundant’ publication by an important scientist that makes the ideas of precursors (of lesser esteem) known.

By the end of the article on the Matthew Effect, Merton gives a short description of these scientists of high esteem. According to him they are often self-confident, they evoke excellence in others and become problem-finders rather than problem-solvers.

Merton then goes on to mention that important scientists tend to be drawn to the same small group of universities (in the USA), where they attract more resources and talented students, leading to a Matthew Effect on institutional level. This observation leads us to the Matthew Effect on a still higher level, namely the Matthew Effect for countries.

The Matthew Effect for countries

The Matthew Effect for Countries (MEC) consists of the systematic deviation in the number of actual (observed) citations from the number of expected citations and this on country level. When studying the underlying mechanism, Bonitz and his collaborators found that almost every journal contributes to the MEC. What actually happens is that each journal ‘redistributes’ citations and this redistribution effect itself can be used as a journal indicator. Only a relatively small group of journals accounts for half of the MEC: these are called the Matthew core journals. Bonitz further notes that the journal Nature has the most Matthew citations. These are citations above the number of expected citations according to the journal’s (standard) impact factor. An important question is “What is the MEC all about? What does it describe?” At first Bonitz proposed three explanations:

- It is just that the rich get richer and the poor get poorer, like it is everywhere in society.
- It is the result of discrimination of one group of countries versus another one.
- It is the result of the language barrier which favours English-speaking countries.

However, on reflection he dismissed these three possible explanations and after thorough analysis concluded that the MEC characterizes the science of
nations by their overall efficiency in scientific performance. In this way the MEC becomes a tool for science policy, because the MEC points to a waste of national resources. From this point of view and taking the enormous scientific growth of China into account\cite{11-13}, it would be very interesting to know if China is still last among the 44 highly productive countries studied by Bonitz.

**Countervailing processes**

In 1988 Merton wrote a follow-up on his original article\cite{14}. Here he discusses, among many other things, why the important universities (he cites Harvard, Columbia and Rockefeller University) have not garnered all (American) Nobel Prize winners. The point he makes is that countervailing processes put an end to ‘infinite’ growth. One of the reasons he mentions is that one department just cannot have a large group of prime movers. This would make the atmosphere (and the necessary resources) intolerable. Moreover, when talented young stars present themselves they do not want to stay in the shadow of old masters (who have seen their best days) and hence they move to another institute. Basically, like in all human endeavours departments, universities and countries rise, disperse and decline.

**The Matthew effect can be linked to the use of concept symbols and default values.**

Small\cite{15} introduced the notion of cited documents as concept symbols or exemplars. He states that scientists carry with them a repertoire of collective concepts and their corresponding concept symbols. Small refers to these symbols as tools-of-the-trade. A typical example is the idea of bibliographic coupling which is – in mind of most informetricians – associated with the work of Kessler\cite{16}, although he was not really the first to think of this concept (Fano being a precursor) and his article in *American Documentation* was only the official version of work that had been described earlier in a series of reports. Extending this idea from the notion of an article as concept symbol to a group of articles we may say that scientific journals may be considered as symbols for a particular type of work. The journals *Nature* and *Science* may be symbols of work deserving the highest visibility. Similarly, in our field, the journals *Scientometrics* and the *Journal of Informetrics (JOI)* may be symbols for the whole field of informetrics and scientometrics. In this way such journals may be considered as a default value: the journal one cites when discussing important work (*Nature and Science*) or work which is considered to be of value for the informetric community (*Scientometrics* and *JOI*). Recent work on default values, e.g. on web browsers, or in clinical procedures\cite{17} discusses that for most people it is difficult to opt out of the default values. Similarly, articles or journals considered as concept symbols become the hard-to-ignore default citation value. This is exactly the Matthew effect at work: the highly visible and highly cited become even more visible and cited, while the less known disappear completely from the “citation” radar. The same idea may explain why astronomers keep on citing (actually cite more and more) the same small group of so-called core journals in their field\cite{18}.

**A bibliometric analysis of Merton’s article on the Matthew Effect**

On June 22, 2011 we downloaded all citations received by MERT68 from the Web of Science (WoS), including SCI-E, SSCI, A&HCI and Conference Proceedings. At that moment the total number of received citations was 865. We present some general characteristics. Garfield is the person who cited MERT68 the most, while most citations occurred in the journal *Scientometrics*, followed by *JASIS(T)* and most citers came from the USA. More importantly citations to Merton’s article are diffused over 126 JCR subject categories. Information science & library science leads this list, followed by sociology.

**Analysis**

Table 1 gives an overview of some global data as obtained by applying WoS analyse function on the set of records that cite MERT68. We did not apply any correction on these data; hence numbers should be seen as approximations.

**Distribution and growth of received citations**

Figures 1 and 2 show the yearly number of received citations by MERT68 and the cumulative number of citations. Already in its year of publication (1968) this article received five citations. During the period 1968-2003 the numbers of received citations fluctuated but showed an increasing trend. Since 2002 a strong increase in the number of received citations is visible.
Indeed, since then MERT68 received more than fifty percent of its total number of citations. This effect is partly due to the yearly growth of the database itself, but clearly the idea of a Matthew Effect in science (and even outside the sciences) still captures the imagination of scientists. The cumulative number of citations shows roughly an exponential increase (a fitted exponential curve has been added to Fig.2).

Although most citing publications are of the article type (75.6%), 9% come are proceedings papers, 8% from reviews and 4% being editorial material. 

Table 1—Overview of data related to MERT68 citations

<table>
<thead>
<tr>
<th>Citing</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publications</td>
<td>865</td>
</tr>
<tr>
<td>Authors</td>
<td>1290</td>
</tr>
<tr>
<td>Subject areas</td>
<td>126</td>
</tr>
<tr>
<td>Countries or regions</td>
<td>50</td>
</tr>
<tr>
<td>Publication sources</td>
<td>425</td>
</tr>
<tr>
<td>Institutions</td>
<td>585</td>
</tr>
</tbody>
</table>

Fig. 1—Yearly citations

Fig. 2—Cumulative citations and best fitting exponential function ($R^2=0.89$)
Next we have a better look at the citing authors, countries and subject areas. In each case we determine a corresponding h-index\textsuperscript{19}. Table 2 gives the eleven authors that cite MERT68 at least six times. Eugene Garfield leads this list; the corresponding citing authors h-index is equal to 7 while its R-index, defined as the square root of the sum of all values in the h-core\textsuperscript{20} is 8.25.

Table 3 shows the most-citing countries or regions (according to the WoS definition of a country). We only show those countries that cite MERT68 at least ten times. Not surprisingly the United States tops this list, followed by the Netherlands. The fact that the Netherlands is number two and that also Belgium figures in this list illustrates the strength of these countries in the field of informetrics. Somewhat surprisingly India does not occur among the 14 most-citing countries. The citing countries h-index is equal to 13 while the R-index is 28.43.

The next table (Table 4) shows a similar list for WoS source publications. In this case we brought the number of references given in \textit{JASIS} and \textit{JASIS(T)} together. This journal list is led by the journal \textit{Scientometrics}, followed by \textit{JASIS(T)}. We note that the recently launched \textit{Journal of Informetrics} is already among the top 10. This list shows that MERT68 is heavily used in the field of information and library science. Consequently, it is no surprise that, in 1995, Robert Merton received the Derek J. de Solla Price award. The citing publications h-index is equal to 10, while the corresponding R-index is 15.36.

Finally, we draw a list for subject areas (as defined in the WoS). Here, information and library science is the number one and other subject areas include engineering sciences, social sciences, humanities and even the so-called multidisciplinary sciences (\textit{Nature} & \textit{Science}). The subject area h-index is equal to 16; the corresponding R-index is 31.67. As WoS subject areas overlap one specific citation may result in two or more citing subject areas. As such the list shown in Table 5 gives just a rough indication of the relative importance of involved subject areas.

We also mention that among the institutions the University of Michigan tops the list followed by Cornell, Columbia and Harvard University. The first nine institutions are all American.

**Diffusion**

The term \textit{diffusion} in the information sciences refers to how scientific ideas are spread from one field to other ones, from one scientist to other scientists, from one country to other countries. Citations of documents are usually considered to be an indication of the diffusion of the ideas, or some of the ideas, put forward in the cited document\textsuperscript{21-23}. In this context the cited documents are the source of diffusion, while the citing documents are the target of diffusion\textsuperscript{24}. Here we consider MERT68 as the source of diffusion and the citation distribution is the diffusion of MERT68’s ideas in science over time. Scharnhorst and Garfield\textsuperscript{1} use MERT68 as an example for a study about the diffusion of scientific ideas.

**Diffusion over countries**

We first consider the diffusion – by citations – over countries. Figure 3 shows that the cumulative diffusion over countries has been increasing linearly.
At point of collecting data for this article, MERT68 had diffused over 50 countries.

Using the terminology introduced by Liu & Rousseau\(^2\) this means that MERT68’s country diffusion breadth was 50. We note that it is possible that one citing article resulting from the collaboration of scientists from different countries, may add several new citing countries, and hence have an important influence on citation diffusion over countries. We did not investigate this. The country diachronous relative diffusion indicator\(^2\) in the year Y+n of a document (or group of documents) published in the year Y defined as:

\[
RD_{n}(O) = \frac{\sum_{j=0}^{n} U(O' + j, Y)}{\sum_{j=0}^{n} CIT(O' + j, Y)}
\]

where U(Y+j,Y) denotes the number of new countries citing at least one of the documents (here this is MERT68 as we study only one document; Y=1968) in the year Y+j. the symbol CIT(Y+j,Y) denotes the number of citations received in the year Y+j by MERT68. The corresponding diffusion curve is

### Table 4—Most-citing journals

<table>
<thead>
<tr>
<th>Publication</th>
<th>Given number of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientometrics</td>
<td>76</td>
</tr>
<tr>
<td>Journal of the American Society of Information Science and Technology</td>
<td>39</td>
</tr>
<tr>
<td>Research Policy</td>
<td>26</td>
</tr>
<tr>
<td>American Sociological Review</td>
<td>19</td>
</tr>
<tr>
<td>Social Studies in Science</td>
<td>17</td>
</tr>
<tr>
<td>American Journal of Sociology</td>
<td>15</td>
</tr>
<tr>
<td>Journal of Informetrics</td>
<td>13</td>
</tr>
<tr>
<td>Current Contents</td>
<td>11</td>
</tr>
<tr>
<td>Annual Review of Information Science and Technology</td>
<td>10</td>
</tr>
<tr>
<td>Organization Science</td>
<td>10</td>
</tr>
</tbody>
</table>
shown in Figure 4. It decreases fast (1971 is the lowest point), but then increases somewhat till the year 1976 and from then on decreases slowly.

The corresponding average diffusion speed\(^{27}\) through the year \(Y+j\) is defined as:

\[
\frac{\sum_{k=0}^{j} Y(k)}{j + 1}
\]

(2)

Its values are shown in Figure 5.

**Diffusion over subject areas**

We recall that MERT68 was published in *Science* which belongs to the area of multidisciplinary sciences. The yearly diffusion over (new) subject areas was rather irregular. Yet, the cumulative number of subject areas in which citations to MERT68 occurred is largely linear, see Fig. 6 (\(R = 0.99\)). Again we note that this growth is not only determined by a larger diffusion of the knowledge contained in Merton’s article, but also by the growth of the database and by the inclusion of new subject areas.

<table>
<thead>
<tr>
<th>Subject area</th>
<th>Given number of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Information Science &amp; Library Science</td>
<td>201</td>
</tr>
<tr>
<td>2 Management</td>
<td>120</td>
</tr>
<tr>
<td>3 Sociology</td>
<td>114</td>
</tr>
<tr>
<td>4 Computer Science, Interdisciplinary Applications</td>
<td>80</td>
</tr>
<tr>
<td>5 Computer Science, Information Systems</td>
<td>72</td>
</tr>
<tr>
<td>6 Social Sciences, Interdisciplinary</td>
<td>58</td>
</tr>
<tr>
<td>7 Business</td>
<td>57</td>
</tr>
<tr>
<td>8 Education &amp; Educational Research</td>
<td>56</td>
</tr>
<tr>
<td>9 Multidisciplinary Sciences</td>
<td>54</td>
</tr>
<tr>
<td>10 Economics</td>
<td>43</td>
</tr>
<tr>
<td>11 History &amp; Philosophy of Science</td>
<td>35</td>
</tr>
<tr>
<td>12 Planning &amp; Development</td>
<td>31</td>
</tr>
<tr>
<td>13 Psychology, Multidisciplinary</td>
<td>31</td>
</tr>
<tr>
<td>14 Psychology, Social</td>
<td>18</td>
</tr>
<tr>
<td>15 Public, Environmental &amp; Occupational Health</td>
<td>17</td>
</tr>
<tr>
<td>16 Biology</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 5—Most-citing subject areas

Fig. 4—Time series of the relative diffusion indicator over countries
Another caveat is again the fact that WoS subject areas overlap so that one specific citation may result in two or more new citing subject areas. As for the diffusion over countries one can also define a relative diffusion index over subject areas, and the corresponding average speed. We leave this for further investigations.

Other effects

Besides the Matthew Effect, other derived or similar effects have been identified in the literature. These effects include the Podunk Effect, the Knudop Effect and the Wehttam Effect. What do these terms mean and who proposed them first?

The word Podunk refers in American English to a small, unimportant and isolated town. The term Podunk university is derived from this and refers to a small-town, unimportant university. The term “Podunk Effect” is a kind of corollary to the Matthew Effect. It is used to describe the fact that scientists from a university or institute of low reputation ‘inherit’ this low prestige. Who first used this term? A search, with the help of a group of friends, called the Knudop Search Group, led to the discovery of Jerry
Gaston’s book *The reward system in British and American science*. On page 121 in the book, Gaston discusses the reward system in science and notes that the system is not perfect. Deviations from perfection require an explanation leading to the Matthew Effect and a new effect for those with lower than expected recognition. He writes:

*The explanation for those with low recognition tends to focus on a process that has not been labeled, and I shall do so here. It is the Podunk Effect.*

Yet, Gaston did not stop there. On page 131 he notes that it should be investigated if, and how often, the opposite of the Matthew Effect and of the Podunk Effect occur. He baptizes these effects the Knudop Effect and Wehttam Effect (even suggesting the pronunciation: wet-tem).

*The Knudop Effect is the process whereby scientists at low-prestige institutions receive more recognition than they apparently deserve. If there is a process opposite from the Podunk Effect, then one must think of the scientists at high-prestige institutions who fail to receive recognition … Let us call it the Wehttam Effect …*

A search for “knudop OR wehttam” in the WoS led to no results. A search for Podunk gave seven results, only one of them being in the field of information and library science, namely Meyer’s article in the *Library Journal*. Yet, this does not mean that no one has studied or at least mentioned these effects. We found several mentions of the Knudop effect in Google Scholar. Liang and Liu found proof of a Knudop effect in the Chinese journal *Acta Physica Sinica*. Yet, these other effects still deserve closer study.

Wu & Wolfram recently proposed four new effects: the self-confidence effect, the Narcissus effect, the other-confidence effect and the flattery effect. Their names clearly refer to a psychological approach and we refer the reader to Wu and Wolfram for further details.

### Matilda Effect

The Matilda Effect, named after the American suffragist Matilda J. Gage of New York, is somewhat similar to the (second part of the) Matthew effect. This effect described by Margaret Rossiter deals with the systematic underrepresentation of women and the reduced recognition of women’s academic contributions. She provides examples of women who have unjustly been denied important scientific prizes including the Nobel Prize, such as Nettie Stevens (chromosomal nature of sex determination), Frieda Robbscheidt-Robbins (haemoglobin regeneration and its relation with diets), Rosalind Franklin (X-ray diffraction images of DNA), Lise Meitner (nuclear fission) and Chien-Shiung Wu (contradiction of the law of conservation of parity). Hence the Matilda effect focuses on the second half of the biblical reference cited by Merton. Rossiter’s article is cited 31 times (in the WoS) up till now and is diffused over 20 WoS subject categories and 16 citing countries. So, in addition to discrimination based on institutional affiliation, (Podunk Effect and Wehttam Effect) reduced recognition may arise from other factors such as gender (women) or geographic region (scientists from developing countries).

### Success versus recognition of research work done by women

Susan Greenfield, the first woman director of the Royal Institution, “concludes” by stating that “Just as science needs more people from diverse cultural backgrounds, it needs more women”. Although women have always done scientific work they have rarely received the recognition they deserved. Consequently, women are often absent from textbooks so that young women have no historical role models and few undergraduates nowadays know Hypatia, Émilie du Châtelet, Maria Agnesi or Mary Somerville. Maybe only Marie Curie has the status of an icon. Political and religious backgrounds, (un)supportive family, math and science “anxiety” all play a role in the underrepresentation of women in science and technology. It is also said that women are less interested than males in seeking the exposure and recognition that accrues from high profile publications or positions.

### Are there physiological differences between male and female brains?

According to some recent studies, it is found that while male and female brains have equivalent potential on broad measures of cognitive ability, there are gender based differences in the neuroanatomical structure of brains such that women’s brains have more white matter and men’s more gray matter. Whereas gray matter represents information...
processing centres in the brain, white matter represents the networking of – or connections between – these processing centres. Based on such scientific studies as well as popular myths, it is often argued that women are more oriented toward and better at assimilating diverse forms of information whereas men prefer to isolate explanations and excel in tasks requiring more local processing. As studying the role of women in science is not the main topic of this contribution we end by mentioning a recent book on this topic. It is written by Stanford sociologist Cecilia Ridgeway and is titled “Framed by Gender: How Gender Inequality Persists in the Modern World.”

Conclusion

It is shown that Merton’s famous article on the Matthew effect has played an important role in the information sciences and beyond. Growth of its received citations continues to grow as does its diffusion over subject areas.

Also derived and related effects such as the Podunk and the Matilda effect have been studied. From this we conclude that the derived effects did not receive a lot of attention and hence we suggest them as a topic for further studies. Our article ends with a short discussion of the underrepresentation of women in science.

Acknowledgments

It is a pleasure to thank Raf Guns (Antwerp University) for correcting some mistakes in an earlier version. Rousseau’s work related to diffusion has been supported by the NSFC grant 71173154.

References