



City Research Online

City, University of London Institutional Repository

Citation: MacFarlane, A., Albrair, A., Marshall, C. R. & Buchanan, G. (2012). Phonological working memory impacts on information searching: an investigation of dyslexia. In: Proceedings of the 4th Information Interaction in Context Symposium. (pp. 27-34). New York, USA: ACM. ISBN 9781450312820 doi: 10.1145/2362724.2362734

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/5217/>

Link to published version: <https://doi.org/10.1145/2362724.2362734>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Phonological working memory impact on Information searching: An investigation of dyslexia

ABSTRACT

A key aspect of searching is the ability of users to absorb information from documents read in order to resolve their ask. One group of users who have problems with reading are dyslexic users, who due to underlying cognitive impairments in phonological processing and working memory, tend to read more slowly and make reading errors. The purpose of this study is to examine the impact of the dyslexia cognitive profile on information searching. Searches were logged for 8 dyslexic and 8 non-dyslexic university students, in order to examine the differences in searching behavior between the two groups. A set of literacy and phonological working memory tasks were also completed, in order to investigate the relationship between these cognitive variables and searching behavior. Results show that there is a significant difference between the two groups on the number of documents being judged irrelevant, and that this cannot be explained by a topic effect. Instead, the number of documents judged irrelevant is significantly correlated with a measure of working memory. This key result provides the research community the first real insight into impact of impaired short term memory on information searching.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Query formulation, Relevance feedback, search process. H.3.3 [User/Machine systems]: Human factors, human information processing.

General Terms

Documentation, Experimentation, Human Factors.

Keywords

Disabilities, dyslexia, information searching, search sessions.

1. INTRODUCTION

Dyslexia is a specific learning difficulty, which affects around 10% of the population [1]. A number of efforts have been made recently in various domains to study the impact of this learning difficulty on the lives of those who have dyslexia. Dyslexia for the purposes of this study is defined as “a learning disorder marked by impairment of the ability to recognise and comprehend written words”. Work is beginning to be done in various areas of the impact of dyslexia on those who use IT, including web navigation [2], assisting users with query errors [3] and the effect of font type and line length on browsing [4]. There is significant interest in the HCI domain, but up till now very little interest in the information seeking and searching domain.

In a pilot study by MacFarlane et al [5], a number of different areas for further research into the impact of dyslexia on information retrieval were defined, including the need to examine the relationships between searching behaviour and reading abilities, speed of language processing and working memory. The aim of this study is to build on the work of MacFarlane et al [5] by examining these relationships, in order to identify which aspects of the dyslexia cognitive profile affect searching behaviour.

The paper is organised as follows. In the related work section we provide a précis of the study carried out by MacFarlane et al [5], and provide an indication of the limitations of that study which this one attempts to address. We also provide some background on the reading literature for adults. We then provide an overview of the methodology used, which is largely identical to the MacFarlane et al [5] study in terms of the procedure and logging mechanism used. However in this study we collect a range of reading-related psychological measures for all participants. Results from logging and dyslexia data, and correlations with reading scores are then examined. The discussion section draws the threads together from this and the previous study to provide more information on what needs to be done in the field. A conclusion is provided at the end.

2. RELATED WORK

The study carried out by MacFarlane et al [5] demonstrated that there were differences between control and dyslexic users when searching on two TREC topics using the Okapi system. Due to the size of the cohort (10 in all, 5 for each group), no significant differences could be established however using qualitative information gathered we hypothesised that various information such as experience etc. may not be able to account for the differences between the two groups. The study found that users who judged more documents as being relevant also tended to view more documents as a whole – irrespective of group membership or topic. In terms of session variables it was hypothesised that there may be an important difference between the groups on the number of searches carried out in a session, and that this variable may not be topic dependent. Evidence of query variables was inconclusive.

From this work a research agenda was established. A much more query interactive data collection method is needed in order to establish the differences between the groups on query variables. Prior data on dyslexia variables such as reading abilities speed and language

processing and working memory is required to better differentiate between the control and dyslexia user groups. This data was not available for the study in MacFarlane et al [5]. Topic effect on search interaction must be taken into account to ensure that any differences between user groups is not effected by the information need. The impact of relevance assessment by both groups is required in order to understand their differences in search and impact on results (precision and recall).

3. THE DYSLEXIA COGNITIVE PROFILE AND INFORMATION RETRIEVAL

Although dyslexia is a heterogeneous disorder, it affects both reading and spelling, and there is compelling evidence that the underlying cognitive deficit lies in impaired phonological processing skills [1]. Phonological processing skills include phonological working memory, which is the ability to hold words for a period of a few seconds whilst processing them, and which is essential for mapping the printed word to its meaning and thereby making sense of text. Reading, spelling and phonological working memory difficulties are expected to impact on the accuracy and speed of information retrieval as follows:

- **Reading difficulties:** Both recognising words (decoding) and understanding them (comprehension) is difficult for dyslexic people, particularly when there is a large amount of text with many rare words, including words that have not previously been encountered. Inaccurate reading is expected to lead to dyslexic people having problems identifying useful terms from documents, and hence identifying concepts to be used to refine their searches. They might even have difficulties in formulating a useful query. Dyslexic readers are not only inaccurate, but also slow, and this slowness is expected to impact on the speed and quality of information retrieval.
- **Spelling difficulties:** If the dyslexic person misspells words, this will cause difficulties in formulating a query. Slow spelling is expected to impact on the speed of information retrieval.
- **Phonological working memory difficulties:** Poor phonological working memory is expected to have a significant impact on information retrieval because if a dyslexic person has difficulty retaining the necessary information in one part (or iteration) of a search session, the implication is that they won't be effective searchers.

In this study we administered tests of reading, spelling and phonological working memory in order to assess their impact on searching behaviour.

4. METHODOLOGY

4.1 Study participants

A total of 22 participants were recruited for the study. Unfortunately the searches for 4 users were not recorded, which left us with a total of 18 useable sets of searches to work with. Of these 18 users, 2 volunteered for the control group but were found to have potential learning difficulties and their data were discarded. These participants were advised accordingly. This left us with a total of 16 users for the study, equally distributed between the control and dyslexia groups. All of these participants were students, 14 language and communication students, 1 law student and 1 economics student. A total of 6 students were studying at various stages for their BSc, and 10 were postgraduates. Undergraduates and postgraduates were equally distributed between groups. However there were many more women than men in the students recruited with only 3 men who took part in our study. Age ranges for both groups were broadly similar, with the control group slightly older. Details of the participants can be found in table 1.

Data	Control Users	Dyslexic Users
<i>Male</i>	1	2
<i>Female</i>	7	6
<i>BSc students</i>	4	2
<i>PG students</i>	4	6
<i>Age Range (years)</i>	20 - 36	20 - 29

Table 1. Details of participants

4.2 Assessments of reading, spelling and phonological working memory

Unlike MacFarlane et al [5] we assessed all our participants on a range of cognitive assessments that are widely used with dyslexic people. Specifically, we carried out the following cognitive tests in order to determine participants' reading, spelling, and phonological working memory skills.

4.2.1 Wide Range Achievement Test-4

The following subtests were used [6].

Reading. This subtest measures reading accuracy. The participant is asked to read a list of words of increasing difficulty, and the examiner scores each word as being either correct or incorrect.

Reading comprehension. This subtest measures the understanding of written languages. The participant silently reads a series of sentences, each of which has a word missing, and has to say aloud a word that fits the meaning of the sentence. The examiner scores each word as being either correct or incorrect for that particular sentence.

Spelling. This subtest measures spelling accuracy. The participant is asked to spell a list of words of increasing difficulty, and the examiner scores each word as being either correct or incorrect.

4.2.2 York Adult Assessment of Dyslexia

The following subtests were used [7]:

Nonsense passage reading. This subtest measures accuracy and speed for reading unknown words. The participant is asked to read aloud two passages, which contain a number of made-up, nonsense words. The passages are scored for the number of nonwords read correctly (according to the letter-sound correspondence rules of English orthography) and for the total amount of time taken to read the passages

Writing speed. This subtest measures speed for writing known words. The participant is asked to write out a particular sentence as many times as they can in two minutes. The total number of words written in that time is scored.

1) Documents

- * Total documents read.
- * Number of Documents judged relevant.
- * Number of Documents judged non-relevant.
- * Number of Documents examined per iteration.
- * Total changes of judgements from relevant to non-relevant.
- * Level of agreement with TREC relevance judgements (%)

2) Sessions

- * Session length (secs).
- * Number of searches (or search iterations).
- * Number of expansions.
- * Number of Hit-lists examined per iteration.
- * Number of Pool views (Documents judged relevant) per iteration.

Figure 1. Log information recorded in database (quantitative variables for study)

4.2.3 Wechsler Adult Intelligence Scales-3

The following subtests were used [8].

Digit span. This subtest measures phonological working memory. In the first part, the examiner reads aloud a series of sequences of digits (from the numbers 1-9, in random order) and the participant repeats each sequence. The sequences increase in length until the participant is no longer able to repeat them correctly. In the second part of the task, the examiner reads aloud a series of sequences of digits that the participant has to repeat in reverse order, i.e. backwards. The total number of sequences recalled correctly, forwards plus backwards, is scored

4.2.4 Further tests

We also asked participants to complete the British Dyslexia Association's Adult Checklist [9], which requires participants to note different diagnostic features of dyslexia that impact on everyday life, for example making mistakes when taking down telephone messages, re-reading paragraphs in order to understand them, and confusing the names of objects. This checklist also yields a raw score that was used in the analysis. All of the dyslexic participants completed the checklist, but only 4 of the 8 control participants.

4.3 Log variables

The logging mechanism used was identical to the MacFarlane et al [5] study. User behavior is stored from the moment the user logs on to the system until they log off. The variables recorded in logging are provided in figure 1. Of primary interest for this study are the document

variables, but session variables are collected for comparison with the prior study. Query variables that were used in the previous study are ignored here

4.4 Topics, collection and system used

Topic No	Control users	Dyslexic users	Total
427	4	5	9
442	4	3	7
Total	8	8	16

Table 2. Topic descriptions and facet analysis

The topics, collection and system used for the experiments was identical to the MacFarlane et al [5] study. Disk 4 and 5 of the TIPSTER collection from TREC 7 and 8 was indexed by Okapi and used for searching. This collection consists of documents from the Financial Times, Federal Register, Foreign Broadcast Information Service and Los Angeles Times. The collection of documents is made up largely of news articles together with US Government documents. We use only two of the general topics from the test collection, namely 427 and 442 - the number of subjects in each group per topic is provided in table 2. Topic 427 focuses on damage to eyes from UV light, while topic 442 focuses on heroic acts. Users were asked to search for one topic only, fewer than many other interactive studies, e.g. Vakkari et al [10] asked participants to search for four topics. Whilst this is problematic (topic has an effect), there are valid ethical reasons for restricting the workload on dyslexic users. This was the strategy used in the MacFarlane et al [5] study. An extra level of analysis is carried out to measure any topic effect.

Topics were assigned to users on a round robin basis. With the control users are equally distributed over the two topics, with dyslexic users the bias is towards topic 427. This because we had to remove 6 participants from the study. This means that topic 427 has more searchers, however there are still sufficient numbers for a topic analysis to be carried out.

4.5 Process and data analysis method

Prior to the experiment a series of cognitive tests was carried out on each participant. On entering the Interaction Lab, the participant was asked to take a seat in front of the participant PC and monitor. The study was explained to the participant, including an overview of the interface, the purpose of the study and the nature of the search task to be conducted in the session. The information sheet was administered, followed by the consent form. The interface was then demonstrated by the experimenter. The procedure for logging in to the IR system, searching using Query terms, modifying Query Terms and judging the relevance of documents was explained until the participant indicated understanding of this. Users completed this procedure only once – no retests were undertaken. Pre and post search surveys were carried out, and the results from these will be reported separately.

Measure	Dyslexic group		Control group		Significance
	Mean (SD)	Range	Mean (SD)	Range	
<i>WRAT word reading subtest</i>	58.25 (3.28)	53-62	64.50 (4.21)	58-69	0.009
<i>WRAT spelling subtest</i>	45.38 (2.13)	43-49	50.13 (3.68)	45-56	0.010
<i>WRAT reading comprehension subtest</i>	49.50 (1.07)	47-50	48.63 (1.69)	46-50	0.328
<i>YAAD non-word passage subtest: time</i>	71.88 (12.18)	57-89	53.63 (25.23)	34-114	0.012
<i>YAAD non-word passage subtest: errors</i>	11.00 (4.69)	3-15	3.00 (2.00)	1-6	0.002
<i>YAAD writing speed subtest</i>	25.86 (4.57)	18-32	30.50 (3.42)	26-35.50	0.058

<i>WAIS digit span</i>	14.25 (3.73)	11-21	21.75 (4.20)	15-26	0.060
<i>Adult checklist</i>	57.00 (5.66)	49-69	33.75 (6.18)	27-42	0.060

Table 3. Dyslexia data results

(***bold**=significant [<0.05], **bold/underline**=very significant [<0.01])

5. RESULTS

5.1 Results of the assessment of reading, spelling and phonological working memory

The two groups differ on several of the cognitive measures (see table 3). The dyslexic group are significantly less accurate than the control group on the single word reading and non-word passage reading tasks, and are significantly slower at reading the non-word passages. They also perform more poorly on the spelling subtest and the writing speed subtest, although the latter just misses significance. However, the dyslexic group score near ceiling on the reading comprehension subtest of the WRAT (the maximum score on this subtest is 51) with a numerically higher mean score than the controls, indicating that despite their reading difficulties, they are able to gain meaning from text just as well as the controls. Finally, the dyslexic group performs marginally worse on the digit span task and reports marginally more dyslexia-type difficulties on the Adult Checklist. There is no doubt then that the dyslexic group does indeed consist of individuals with the characteristic dyslexia cognitive profile, despite the fact that they are all able to cope with the literacy demands of studying at university.

5.2 Search Log Results

One issue stands out in the data (see table 4), the difference between the controls and dyslexic users on total documents read is very nearly significant with a substantial difference in the average number of documents read of 42 per dyslexic user (with control users reading many more – on average 56.8 documents). Virtually all of the difference between the two groups is down to the number of documents being judged irrelevant with an average of 14 more documents being judged irrelevant by control users – this difference is significant. The number of documents judged as being relevant is virtually the same for both groups. The level of agreement with TREC relevance assessments is also virtually the same (however is very low for both groups – this may well be a factor of the participants disciplines). On average control users are reading 17.5 documents per iteration, which is qualitatively different, but is not statistically significant. In terms of session data, there are no significant differences, although it is noteworthy that the number of searches carried out by dyslexic users is very nearly double that of controls. This is the reverse of what was found in the previous study [5]. Dyslexic users in this study tended to do more within an iteration than controls. This contrasts rather starkly with the previous study [5] where dyslexic users viewed slightly more documents per iteration. In fact the pattern of reading and relevance assessment is very different in nature than the previous study. For example control readers judged more documents as being irrelevant compared with our previous study where dyslexic users were judging more documents as being irrelevant. We posit a significant reason for the differences between the results in this study and the previous one. The participants in the previous study were very IT literate and had significant experience of computing, some to PhD level. Students in this study were mostly from the social sciences without the same IT background. The cohort in the previous study had clearly built up strategies which the cohort in this study have not yet been able to do – interestingly this appears to have had a bigger impact on the control user group than the dyslexic user group. In general the log results for dyslexic users in both studies is very similar.

Measure	D Avg.	D SD	C Avg.	C SD	Significance
<i>Documents</i>					
<i>Total Documents read</i>	42.0	10.5	56.8	16.9	0.05
<i>Documents judged relevant</i>	14.9 (35%)*	7.9	15.6 (27%)*	6.8	0.84
<i>Documents judged non-relevant</i>	27.1 (65%)*	12.1	41.1 (73%)*	13.3	0.04
<i>Documents examined per iteration</i>	6.9	4.2	17.5	15.0	0.08
<i>Changes of judgments from relevant to non-relevant</i>	0.6	0.9	1.1	1.5	0.43
<i>Level of agreement with TREC relevance judgments</i>	14.9%	7.9%	15.6%	6.8%	0.84
<i>Sessions</i>					
<i>Session length (secs)</i>	2020	270.6	2025	474.1	0.98
<i>Number of searches</i>	9.1	7.2	5.6	4.4	0.26
<i>Hit-lists examined per iteration</i>	20.5	13.3	11.4	9.3	0.13

<i>Pool views (Documents judged relevant) per iteration</i>	2.9	2.2	3.5	3.4	0.67
---	-----	-----	-----	-----	------

Table 4. Search log result averages and standard deviations (Dyslexic vs. Control users).

(*Percentage of total. **bold**=significant [>0.05], **bold/underline**=very significant [>0.01])

Measure	427 Avg.	427 SD	442 Avg.	442 SD	Significance
<i>Documents</i>					
<i>Total Documents examined</i>	47.3	6.7	51.0	20.4	0.65
<i>Documents judged relevant</i>	14.7	8.0	15.7	6.8	0.80
<i>Documents judged non-relevant</i>	32.57	10.18	35.33	17.31	0.71
<i>Documents examined per iteration</i>	14.24	13.94	10.22	4.6	0.47
<i>Changes of judgments from relevant to non-relevant</i>	1.09	1.30	0.67	1.41	0.45
<i>Level of agreement with TREC relevance judgments</i>	14.71%	8.04%	15.7%	6.8%	0.80
<i>Sessions</i>					
<i>Session length (secs)</i>	1811	401.7	2187	362.8	0.07
<i>Number of searches</i>	9.7	8.6	5.6	2.2	0.18
<i>Hit-lists examined per iteration</i>	15.1	13.5	16.6	11.6	0.83
<i>Pool views (Documents judged relevant) per iteration</i>	5.3	3.0	1.6	1.1	<u>0.004</u>

Table 5. Search log result averages and standard deviations (topic 442 vs. topic 427).

(***bold**=significant [>0.05], **bold/underline**=very significant [>0.01])

Examination of the same data (see table 5) using topic as the variable demonstrates that the reading variables such as documents examined and judged relevant/irrelevant are broadly similar in terms of averages and are not significantly different. This provides us with some confidence that in this case there is no significant topic effect, and we can rule out this variable in trying to explain any differences between the two user groups on reading, writing and memory variables. Most other variables on the topic comparison are broadly similar and not significant, although the number of searches in a session for topic 427 was nearly double that of topic 442, and there was around 300 seconds difference in the session lengths between the two topics, but this is well within the range of SD's for both topics. The only variable where a very significant difference is registered is pools views, with users searching on topic 427 on average viewing the list of relevant documents over three times more than topic 442, meaning that these users felt the need to refer back to documents in the latter topic. There is no significance between the user groups, and this factor does not appear to impact on the overall result of this study.

5.3 Correlation of dyslexia and search results

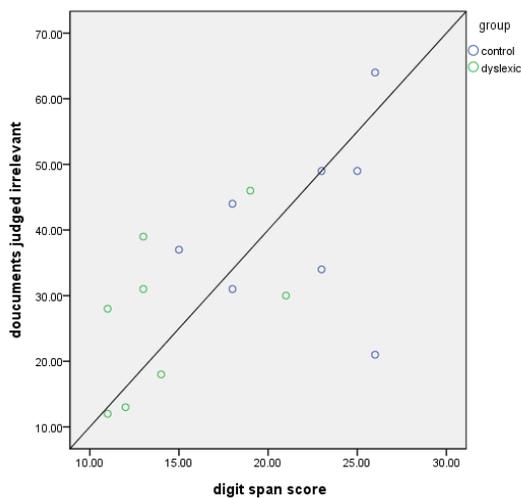


Figure 2. Correlation between number of documents judged irrelevant and digit span score.

Group differences in the total number of documents examined are driven by the number of documents that the two groups judge as irrelevant. In order to investigate which of the underlying cognitive variables can account for group differences in the number of documents judged irrelevant, we carried out a set of bivariate correlations between the number of documents judged irrelevant and the cognitive measures. The only significant correlation was with the measure of phonological working memory, digit span score, $r(16) = 0.586$, $p = 0.017$, see figure 2. This indicates that individuals with better phonological working memory, as indexed by the digit span task, are likely to judge a greater number of documents as irrelevant. One other correlation approached significance, which was that with the British Dyslexia Association's Adult Checklist, $r(12) = -0.506$, $p = 0.093$. If we'd been able to test all the controls on the checklist, then this result might well have reached significance. The negative value of the correlation means that individuals with fewer symptoms of dyslexia in their everyday functioning are likely to judge a greater number of documents as irrelevant.

6. DISCUSSION

The two key results from the research are the significant difference between dyslexic and control users on documents judged irrelevant, and the correlation found on that variable against the digit span working memory variable. Any difference on the documents read between control and dyslexic groups can be explained on documents judged irrelevant. When examining documents judged irrelevant on a topic basis, no significant difference was found, and there is no evidence for a topic effect on the difference on that variable. What do these key results tell us? Or in other words why does memory have an impact on the numbers of documents being judged irrelevant? Why do other variables such as reading do not have such an impact, and no significant difference is found on the dyslexia reading variables?

A key issue is that memory is a higher order cognitive function which drives reading i.e. poor short term memory inhibits reading abilities. This is clearly born out by the evidence presented in this study [needs finessing].

7. CONCLUSION

The key result on the impact of short term memory on information searching is the key result from this research. We have been able to achieve our aim of identifying which aspects of the dyslexia cognitive profile have an impact on search behavior – the impaired short term memory function of the dyslexic user. From this study and the previous one carried out [5], we conclude that close attention needs to be paid to the cohort of participants in any study of this type, in particular the previous knowledge of IT needs to be established as it appears to have a significant impact based on the evidence from the two studies. The searching behavior in both studies carried out is broadly similar, but the control user behavior is very different. This can be explained by the cohort in this study being less IT literature than the previous study (who were all educated to a degree level in the subject). From this we can conclude that intervention in terms of training search is more likely to lead to better results for control users rather than dyslexic users. The next step for us in terms of research is how to establish strategies and tools for helping dyslexic users when searching, given their impaired memory function.

8. ACKNOWLEDGMENTS

The research presented here was part funded by City University London, via a pump prime funding grant. The research methodology used was approved by the University Ethics Committee.

9. REFERENCES

- [1] Snowling, M. 2000. *Dyslexia, 2nd Edition*. Blackwell Publishing.

- [2] Al-Wabil, A. 2009. *The effect of dyslexia on Web Navigation*, PhD thesis, April. City University London.
- [3] Sitborn, L., Bellot, P., Blache, P. 2007. Phonetic based sentence level rewriting of questions type by dyslexic spellers in an information retrieval context. In *Proceedings of Interspeech 2007 – EuroSpeech, Antwerp, Belgium*.
- [4] Ling, J. and van Schaik, P. 2006. The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies*, 64, 395-404. DOI= <http://dx.doi.org/10.1016/j.ijhcs.2005.08.015>
- [5] MacFarlane, A., Al-Wabil, A., Marshall, C.R., Albrair, A., Jones, S.A. and Zaphiris, P. 2010. The effect of dyslexic on Information Retrieval: a pilot study *Journal of Documentation*, 66(3), 307-326. DOI= <http://dx.doi.org/10.1108/00220411011038421>
- [6] Wilkinson, G. S., & Wilkinson, G. J. 2006. *Wide Range Achievement Test, 4th Edition*. Florida, USA: Psychological Assessment Resources, Inc.
- [7] Hatcher, J., & Snowling, M. J. No Date. York Adult Assessment of Dyslexia. Available on: <http://www.york.ac.uk/psychology/research/groups/crl/assessments/>, visited: 26th February 2012.
- [8] Wechsler, D. 1997. *Wechsler Adult Intelligence Test – 3rd Edition*, The Psychological Corporation: San Antonio.
- [9] British Dyslexia Association. No Date. *Adult Checklist* Available on: <http://www.bdadyslexia.org.uk/files/Adult%20Checklist.pdf> . visited: 26th February 2012.
- [10] Vakkari, P., Jones, S., MacFarlane A, and Sormunen E. (2004). Query exhaustivity and Relevance Feedback, *Journal of Documentation*, Vol. 60 (2), 109-127. DOI= <http://dx.doi.org/10.1108/00220410410522016>.