

# **Information-Seeking Behaviour of Physicists and Astronomers: An Intradisciplinary Study**

**Hamid Reza *Jamali Mahmuei***

A thesis submitted in fulfilment of the requirements for the degree  
of Doctor of Philosophy of the University of London



School of Library, Archive and Information Studies  
University College London  
University of London

April 2008

UMI Number: U591583

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U591583

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## Abstract

The study of information-seeking behaviour of scientists has been one of the main concerns of librarians and information scientists since mid twentieth century and yet we need to improve our understanding of their information behaviour in order to maximise the efficiency of information services provided. This thesis studies the information-seeking behaviour of physicists and astronomers with an intradisciplinary approach in order to look at similarities and dissimilarities among the subfields within physics and astronomy. The study also looks at the information-seeking behaviour of people with different academic status and investigates the information-seeking activities of physicists and astronomers in different stages of research projects with the focus of the thesis being research related information-seeking behaviour. Moreover, the research investigates reading behaviour and publishing patterns of physicists and astronomers.

The study is a mixed-methods study that uses both qualitative and quantitative approaches. The population of the study included the staff and PhD students in the Department of Physics and Astronomy of University College London. Fifty-six face-to-face semi-structured interviews were conducted, an online questionnaire survey of 114 respondents (out of 242 sample, 47% response rate) was carried out and 88 information-event cards were completed by participants.

The findings of the study showed that although some similarities exist in information-seeking behaviour of people in the different subfields of physics and astronomy, each subfield has its own characteristics. Variations were found with regard to different aspects of information-seeking behaviour including the reliance on e-print archives and journal articles, methods used for keeping up-to-date and methods used for identifying articles. The study showed the importance of human information sources and informal communication in the information-seeking behaviour of physicists and astronomers and highlighted the need for and the value of looking at narrower subject communities within disciplines for a deeper understanding of the information behaviour of scientists.

## Declaration of originality

I declare that the work presented in the thesis is my own work and that, to the best of my knowledge and belief, is comprised of original research and ideas, except as acknowledged in the text. It contains no material submitted for any other degree at this or any other university.

Signed: Hamid Reza *Jamali Mahmuei*

PhD candidate

## Dedication

*To my family  
For their loving support  
Without which I could never have come this far*

به خانواده عزیزم  
که محبت و تشویقشان همیشه مایه دلگرمی و امیدم بوده است

## Acknowledgement

I would like to begin by expressing my deepest gratitude to all Iranian people, whose precious resources were used through the Ministry of Science, Research and Technology to fund my PhD. I humbly wish that the experience and the knowledge I have gained during my studies in England will make a little contribution in the progress of my beloved homeland, Iran.

Completing doctoral research is a difficult process with so many ups and downs. I am deeply indebted to my erudite supervisor, Professor David Nicholas. To Dave, thank you for your advice that has been the guiding light during the adventurous journey of my PhD. You provided me with the opportunity to become engaged in several research projects, and helped me obtain invaluable experience and contribute to numerous publications. I should also like to thank my second supervisor. To Dr. Rob Miller, thank you for your kind support.

My bottomless appreciation also goes to my friend, Dr. Shant Narsesian. Shant, your friendship made my staying in England cheerful and I learned many things from you. There are a couple of other friends, all research students at the time and doctors now, who made my graduate life not only endurable, but also enjoyable. To Dr. Saeid Asadi, Mr. Shahram Sedghi, Dr. Yazdan Mansourian, Dr. Alireza Norouzi, Dr. Panayiota Polydoratou, Miss Mehrnoush Mozaffarian, Mrs Isabel Galina, Dr. Imad Karam, Dr. Ehsan Hamadani, Dr. Cristina Arciniegas, Mrs. Debora Westwell and Mr. Amir Chahardehi,; thank you for kindly sharing with me your happy and sad moments, your experiences and advice and for offering your help whenever I was in need.

Many thanks also go to a few other friends and scholars who assisted me with their knowledge and advice during my studies, especially Dr. Ian Rowlands, Dr. Jon Rimmer, Professor Carol Tenopir and Dr. Ali Shiri.

I must acknowledge the help of PhD students and the staff in the Department of Physics and Astronomy at University College London that made it possible for me to undertake this research project. Their assistance was beyond my expectation.

There are many other people who have had an impact on my studies so far. Last but not least, I would like to thank them *en masse*. Among them, are the people in the School of Library, Archive and Information Studies at UCL, as well as my previous lecturers in Iran from whom I learned a lot about Library and Information Sciences.

## Table of contents

Abstract .....	1
Declaration of originality .....	2
Dedication .....	3
Acknowledgement.....	4
Table of contents .....	5
List of figures .....	10
List of tables .....	13
CHAPTER 1 - INTRODUCTION .....	14
1.1. Introduction .....	14
1.2. Statement of the problem .....	14
1.3. Motivations for the study .....	15
1.4. Aims and objectives .....	18
1.5. Research questions .....	20
1.6. Scope of the study .....	20
1.7. Distinctive quality of the study .....	23
1.8. Thesis' outline .....	23
1.9. Summary .....	24
CHAPTER 2 - LITERATURE REVIEW .....	25
2.1. Introduction .....	25
2.2. Information behaviour .....	25
2.3. Information-seeking behaviour .....	26
2.4. Approaches to information behaviour research.....	28
2.4.1. Cognitive approach.....	28
2.4.2. Social approach .....	29
2.4.3. Domain-analytic approach .....	31
2.5. Models of information-seeking behaviour .....	33
2.5.1. Ellis's model of information seeking strategies .....	35
2.6. Information-seeking behaviour in academic environment.....	39
2.7. Physicists and astronomers.....	40
2.7.1. Information-seeking behaviour .....	41
2.7.2. Reading behaviour.....	43
2.7.3. Journals, e-prints and open access.....	46
2.7.4. Scholarly Communication and IT .....	49
2.7.5. Students' information-seeking behaviour .....	53
2.8. Summary .....	55
CHAPTER 3 - METHODOLOGY .....	57

3.1.	Introduction .....	57
3.2.	Qualitative approach vs quantitative approach .....	58
3.3.	Qualitative and quantitative research in LIS .....	59
3.4.	Mixed methods .....	60
3.5.	Research design .....	62
3.6.	The research population .....	65
3.6.1.	UCL Department of Physics and Astronomy .....	65
3.7.	Research methods used .....	67
3.7.1.	Desk Research .....	68
3.7.2.	Interview .....	69
3.7.2.1.	Sampling .....	70
3.7.2.2.	Email interviewing .....	71
3.7.2.3.	Pilot interview study .....	72
3.7.2.4.	Interview data collection procedure .....	72
3.7.2.5.	Interview protocol .....	74
3.7.2.6.	Analysing interviews .....	75
3.7.2.7.	Memo writing .....	76
3.7.3.	Information-event card .....	77
3.7.4.	Questionnaire .....	80
3.7.4.1.	Design of online questionnaires .....	81
3.7.4.2.	Sampling and conducting the survey .....	83
3.7.4.3.	The questionnaire's content .....	84
3.7.4.3.1.	Keeping up-to-date (Q 2-4) .....	84
3.7.4.3.2.	Methods used for identifying research articles (Q 5-6) .....	85
3.7.4.3.3.	Reading behaviour (Q 7-11) .....	85
3.7.4.3.4.	Statements (Q 12) .....	86
3.7.4.3.5.	Publishing behaviour (Q 13-16) .....	86
3.7.4.3.6.	Characteristics of research subfield (Q 17 & 18) .....	86
3.7.4.3.7.	Problems (Q 19) .....	87
3.7.4.3.8.	Demographics (Q 20-) .....	87
3.7.4.3.9.	Comments .....	88
3.8.	Validity and reliability .....	88
3.8.1.	Quantitative part of the research .....	88
3.8.1.1.	Validity .....	88
3.8.1.2.	Reliability .....	89
3.8.2.	Qualitative part of the study .....	90
3.8.2.1.	Credibility .....	90
3.8.2.1.1.	Objectivity (neutrality) .....	91
3.8.2.1.2.	Validity (Trustfulness) .....	91
3.8.2.1.3.	Reliability (replicability) .....	92
3.9.	Limitations of the study .....	93
3.10.	Ethical issues .....	94
3.11.	Summary .....	95
CHAPTER 4 - DEMOGRAPHICS OF THE SAMPLE .....		96
4.1.	Introduction .....	96
4.2.	The case study department .....	96



4.3.	Interview participants' profile .....	97
4.4.	Information-event booklets' participants .....	101
4.5.	Survey participants .....	103
4.6.	Summary .....	105
CHAPTER 5 - RESULTS .....		106
5.1.	Introduction .....	106
5.1.1.	Presentation .....	106
5.1.2.	Structure of the results.....	107
5.1.3.	Interweaving the data .....	109
5.2.	Initiating research and seeking information .....	112
5.2.1.	Staff .....	112
5.2.2.	Research Students.....	117
5.3.	Methods used for identifying articles.....	120
5.3.1.	By frequency of use.....	120
5.3.2.	By users' academic status.....	122
5.3.3.	By users' gender .....	123
5.3.4.	By users' type of research .....	124
5.3.5.	By interdisciplinarity of the field .....	125
5.3.6.	By scatter of the field's literature .....	127
5.3.7.	By age of article .....	129
5.3.8.	By users' research group .....	129
5.4.	Methods used for keeping up-to-date.....	130
5.4.1.	By dependency on various methods.....	133
5.4.2.	By importance of keeping up-to-date.....	134
5.4.3.	By age of participants.....	135
5.4.4.	By academic status of participants .....	136
5.4.5.	By type of research.....	137
5.4.6.	By interdisciplinarity of the field .....	138
5.4.7.	By scatter of the literature .....	139
5.4.8.	By research group.....	140
5.4.8.1.	Atmospheric Physics (AP) .....	141
5.4.8.2.	High Energy Physics (HEP).....	142
5.4.8.3.	Condensed Matter and Materials Physics (CMMP).....	143
5.4.8.4.	Astronomy and Astrophysics (AA).....	144
5.4.8.5.	Theoretical Molecular Physics (TMP) .....	145
5.4.8.6.	Atomic, Molecular, Optical and Positron Physics (AMOP) ...	146
5.4.8.7.	Optical Science Laboratory (OSL).....	146
5.4.9.	Role of conferences in information seeking.....	147
5.4.10.	Role of research environment and colleagues in information seeking 149	
5.5.	Problem-specific information seeking.....	154
5.5.1.	Web searching .....	157
5.6.	Accessing information.....	160
5.7.	Changes in information-seeking over time .....	164
5.8.	Reading behaviour.....	167
5.8.1.	Reading quantity.....	167

5.8.2.	Reading and other activities .....	167
5.8.3.	By research group.....	169
5.8.4.	By methods used for identifying articles.....	170
5.8.5.	By importance of keeping up-to-date .....	171
5.8.6.	By age of the material read.....	172
5.8.7.	Reading from screen.....	175
5.8.8.	Role of abstracts .....	176
5.8.9.	Credibility of information .....	177
5.9.	Publishing behaviour .....	179
5.9.1.	Writing approach.....	180
5.9.2.	Publishing approach, journals .....	181
5.9.3.	Publishing approach, e-prints .....	185
5.10.	Problems and difficulties.....	189
5.10.1.	Access to information resources.....	193
5.10.2.	Information overload.....	194
5.10.3.	Technical information and minor pieces of information.....	197
5.10.4.	Inefficient searches.....	198
5.10.5.	Ambiguity of the abstracts .....	203
5.10.6.	Lack of time.....	204
5.10.7.	Personal information management.....	204
5.10.8.	Browsing and serendipity .....	204
5.11.	Summary .....	206
CHAPTER 6 - DISCUSSIONS AND CONCLUSIONS .....		208
6.1.	Introduction .....	208
6.2.	Information-seeking behaviour of physicists and astronomers.....	208
6.2.1.	By subfields.....	208
6.2.1.1.	Astronomy and Astrophysics (AA).....	209
6.2.1.2.	Atomic, Molecular, Optical and Positron Physics (AMOP) ...	209
6.2.1.3.	Atmospheric Physics (AP) .....	210
6.2.1.4.	Condensed Matter and Materials Physics (CMMP).....	210
6.2.1.5.	High Energy Physics (HEP).....	211
6.2.1.6.	Optical Science Laboratory (OSL).....	211
6.2.1.7.	Theoretical Molecular Physics (TMP) .....	212
6.2.2.	By academic status .....	214
6.2.3.	By research stages .....	216
6.3.	Reading behaviour.....	219
6.4.	Publishing behaviour.....	222
6.5.	Problems in information seeking.....	223
6.6.	Contribution of the study.....	225
6.7.	Models of information-seeking behaviour .....	227
6.8.	Further research.....	228
6.9.	A final word.....	229
6.10.	Summary .....	230
BIBLIOGRAPHY .....		232
APPENDICES.....		256

---

Appendix 01: The letter to the Head of the Dept. of Physics and Astronomy .....	256
Appendix 02: The permission from the Head of the Department of Physics and Astronomy for undertaking the research .....	258
Appendix 03: The note published in the weekly email newsletter of the Dept. of Physics and Astronomy .....	259
Appendix 04: The invitation email for participation in the interviews .....	260
Appendix 05: The interview protocol .....	261
Appendix 06: The Information-Event Card .....	262
Appendix 07: The invitation email for participation in the survey .....	263
Appendix 08: The reminder email for participation in the survey .....	264
Appendix 09: The online questionnaire – staff version .....	265
Appendix 10: The online questionnaire - PhD students version .....	271
Appendix 11: Cluster analysis of questionnaire survey results.....	276
Appendix 12: List of publications .....	280

## List of figures

Figure 2.1. A nested model of conceptual areas of information behaviour (Wilson, 1999:263).....	27
Figure 2.2. Wilson’s model of information behaviour (Wilson, 1999:251). ....	34
Figure 2.3. A stage process version of Ellis’s model (Wilson, 1995: 255). ....	36
Figure 2.4. Stages in the information-seeking behavior of academic social scientists (Meho and Tibbo, 2003:584). ....	37
Figure 3.1: Simultaneous equivalent status QUAL/QUAN design of mixed methods research (Creswell, 2003, p. 214). ....	60
Figure 3.2: Sequential equivalent status QUAN/QUAL design of mixed methods research. ....	61
Figure 3.3: Sequential equivalent status QUAL/QUAN design of mixed methods research. ....	61
Figure 3.4: Dominant- less dominant mixed methods research (Creswell, 2003, p. 214). ....	61
Figure 3.5: The stage of the research process. ....	64
Figure 3.6: Number of articles published by UCL in physics and astronomy (ISI data).....	66
Figure 3.7: Number of citations received by articles published by UCL in physics and astronomy (ISI data). ....	66
Figure 3.8: Information-Event Card .....	79
Figure 3.9: Screenshot of the online questionnaire survey, the staff version. ....	82
Figure 4.1: Percentage distribution of interviewees by status. ....	100
Figure 4.2: Percentage distribution of interviewees by type of research. ....	100
Figure 4.3: Percentage distribution of interviewees by research group. ....	101
Figure 5.1: Percentage frequency distribution of methods used for identifying articles. ...	121
Figure 5.2: Percentage frequency distribution of methods used for identifying last read article. ....	122
Figure 5.3: Percentage breakdown of the most used methods for identifying articles by respondents' status. ....	123
Figure 5.4: Percentage breakdown of the most used methods for identifying articles by gender. ....	124
Figure 5.5: Percentage breakdown of the most used methods for identifying articles by type of research of respondents. ....	125
Figure 5.6: Percentage breakdown of use of interdisciplinary literature by research group. ....	125
Figure 5.7: Percentage breakdown of the most used methods for identifying articles by interdisciplinarity of the field. ....	126
Figure 5.8: Percentage breakdown of the most used methods for identifying articles by scatter of subfields’ literature. ....	128
Figure 5.9: Percentage breakdown of the most used methods for identifying articles by scatter of subfields’ literature. ....	128
Figure 5.10: Percentage breakdown of age of last read article by methods used for finding it. ....	129

Figure 5.11: Percentage breakdown of methods used for finding last article read by respondents' research group. ....	130
Figure 5.12: Percentage frequency distribution of importance level of keeping up-to-date. ....	131
Figure 5.13: Percentage breakdown of importance level of keeping up-to-date by age of respondents. ....	132
Figure 5.14: Percentage breakdown of importance level of keeping up-to-date by users' academic status. ....	132
Figure 5.15: Percentage frequency distribution of dependency on various methods for keeping up-to-date. ....	133
Figure 5.16: Percentage frequency distribution of the top three ranked methods for keeping up-to-date. ....	134
Figure 5.17: Percentage breakdown of the most used methods for keeping up-to-date by importance or keeping up-to-date. ....	135
Figure 5.18: Percentage breakdown of top ranked methods for keeping up-to-date by age of respondents. ....	136
Figure 5.19: Percentage breakdown of top ranked methods for keeping up-to-date by respondents' academic status. ....	137
Figure 5.20: Percentage breakdown of top ranked methods for keeping up-to-date by type of research of respondents. ....	138
Figure 5.21: Percentage breakdown of top ranked methods for keeping up-to-date by interdisciplinarity of subfields. ....	139
Figure 5.22: Percentage breakdown of the top ranked methods for keeping up-to-date by scatter of subfields' literature. ....	140
Figure 5.23: Percentage breakdown of the top ranked methods for keeping up-to-date by respondents' research group. ....	141
Figure 5.24: Percentage breakdown of responses to the statement on access to non-online articles by respondents' research group. ....	162
Figure 5.25: Percentage breakdown of responses to the statement on access to non-online articles by respondents' status. ....	163
Figure 5.26: Percentage frequency distribution of respondents by status by number of articles read. ....	168
Figure 5.27: Percentage frequency distribution of respondents by research group by number of preprints read. ....	169
Figure 5.28: Percentage frequency distribution of respondents by research group by number of articles read. ....	170
Figure 5.29: Percentage frequency distribution of respondents by methods used for identifying articles by number of articles read. ....	171
Figure 5.30: Percentage frequency distribution of respondents by importance of keeping up-to-date by number of articles they read. ....	172
Figure 5.31: Percentage frequency distribution of age of last read research article. ....	173
Figure 5.32: Percentage breakdown of age of last read article by age of respondents. ....	173
Figure 5.33: Percentage breakdown of age of last read article by respondents' status. ....	174
Figure 5.34: Percentage breakdown of age of last read article by respondents' research group. ....	175
Figure 5.35: Percentage frequency distribution of respondents by whether they deposit their articles by research groups. ....	187

---

Figure 5.36: Percentage breakdown of depositing behaviour by subgroups of astronomy and astrophysics. ....	188
Figure 5.37: Percentage frequency distribution of respondents by whether they deposit their articles by number of preprints they read in a month. ....	189
Figure 5.38: Percentage frequency distribution of problems. ....	192
Figure 5.39: Frequency distribution of problems by respondents' research group. ....	192
Figure 5.40: Percentage breakdown of responses to the statement on subject keyword searching by respondents' research group. ....	202
Figure 5.41: Percentage breakdown of responses to the statement on subject keyword searching by respondents' status. ....	203
Figure 6.1: Variations in information-seeking behaviour based on academic status. ....	216
Figure 6.2: Stages of information-seeking in a research project. ....	219

## List of tables

Table 3.1: Differences between quantitative and qualitative approaches based on Hammersley (1992) and Punch (2005).....	58
Table 3.2: Research groups in the Department of Physics and Astronomy.....	67
Table 3.3: Research Methods applied in the study. ....	68
Table 4.1: Distribution of population by gender and status. ....	97
Table 4.2: The profile of the participants and the interview sessions – research students....	98
Table 4.3: The profile of the participants and the interview sessions – academic staff.....	99
Table 4.4: The profile of the participants who completed information-event cards. ....	102
Table 4.5: Number of completed information-event cards by participants’ status and gender. ....	103
Table 4.6: The size of the sample and the response rate of the survey.....	103
Table 4.7: Distribution of the respondents by gender.....	103
Table 4.8: Distribution of the respondents by age. ....	104
Table 4.9: Distribution of the respondents by status.....	104
Table 4.10: Distribution of the respondents by type of research. ....	105
Table 4.11: Distribution of the respondents by research group. ....	105
Table 5.1: Methods used for investigating different issues. ....	111
Table 5.2: Distribution of the information seeking events by their outcome.....	155
Table 5.3: Distribution of information seeking events by type of information sought.....	156
Table 5.4: Distribution of resources and methods used in information seeking events.....	157
Table 5.5: Distribution of respondents by their level of agreement with the statement: ‘If an article is not available online, it’s probably not worth the effort to obtain it’.....	161
Table 5.6: Number of articles and preprints read. ....	167
Table 5.7: Number of articles and preprints read by academic status. ....	168
Table 5.8: Number of journal articles and preprints read by research group.....	170
Table 5.9: Distribution of respondents by whether they deposit their articles in e-print archives.....	185
Table 5.10: Time of depositing articles in e-print archives. ....	185
Table 5.11: Reasons why not depositing articles in e-print archives.....	186
Table 5.12: List of problems and difficulties of participants in their information seeking activities.....	191
Table 5.13: Distribution of respondents by their level of agreement with the statement: ‘I tend to avoid using subject keywords and phrases when searching databases to find articles because it brings up too many results’.....	201
Table 6.1: Information-seeking characteristics of research groups. ....	213

## CHAPTER 1 - INTRODUCTION

### 1.1. Introduction

In the current information age, seeking information is a fundamental function and will continue to be so. Generating, storing, collecting, organising, seeking, searching and retrieving information are still among the main concerns of this age. Academic scientists are an important ring in the chain of generating and using scholarly information. New technologies have affected every function in universities, libraries, and information services. Scientists now more and more work in an electronic information-rich environment and physicists and astronomers are among pioneer scientists to take advantage of new information technologies.

This thesis is a study of the information-seeking behaviour of physicists and astronomers. The research utilizes mixed-methods to study the information-seeking behaviour of physicists and astronomers and to derive their behavioural patterns. Interviews, a questionnaire survey and critical incident technique were used to explore information behaviour of academic physicists and astronomers, with a particular interest in intradisciplinary comparison. The population studied in this research was taken from the Department of Physics and Astronomy at the University College London (UCL).

This introductory chapter aims to outline the research, delineate the research problem and explain the aims and objectives of the research. Moreover, the chapter seeks to illustrate the scope of the study which is physics and astronomy.

### 1.2. Statement of the problem

How do scientists really discover, select and use the countless information and communications resources available to them? This question has been addressed through questionnaire surveys, interviews and log analyses for a number of years, with varying degrees of success in identifying characteristic habits and behaviour. Studying the information behaviour of scientists has been one of the main concerns of librarians and information scientists at least since The Royal Society Scientific Information Conference of 1948 (Royal Society, 1948). There have been many studies which have investigated different aspects of the information behaviour of scholars with different methodologies, objectives and from different perspectives (see Chapter Two – Literature Review). Many of these studies have investigated specific aspects of the information-seeking behaviour of large groups of scientists, thanks to which we have developed a reasonably good understanding of the scholarly communications



and information-seeking behaviour of the scholars. However, this is by no means an indication that there is no need for further study on information behaviour of scholars. As information technologies, which nowadays are major means of information service provision, develop, information services are improved and as a result information seeking activities of scientists go through changes and adjustments. This is a cycle where research on information behaviour of scholars leads to better information services and improved information services might make the scholars alter their information seeking activities and behaviour, hence the need for continuous study of the information-seeking behaviour of scholars.

The idea for this research initiated from the importance of studying and understanding the information-seeking behaviour of scholars. Scholars are key players in Scientific, Technical and Medical (STM) publishing, which is a multibillion dollar industry with more than 12 billion dollars predicted revenue for 2005 (Jastrow, 2004). Tenopir and King estimated that the entire scientific scholarly journal system in the United States of America (USA) expends about \$45 billion in a year. The majority of these expenses, which does not include exchange of money such as subscriptions, cover scientists' time and other resources associated with authorship (9% of the total) and reading (78% of the total) (2000:4). As Tenopir and King stated:

scientists' time is a critical resource and any system innovations and service decisions should take into serious consideration the consequences on this time, for example, by minimizing the authors' time and readers' time spent required to identify, locate and acquire the information and time spent reading and assimilating the information. (2000:4)

A more efficient scientific information system could save a knowledge-driven country such as the USA millions of dollars. Physics and astronomy are known as expensive sciences. Nowadays, conducting research in certain areas of physics and astronomy is not feasible for countries unless they are done as multinational collaborative projects. The financial factors and the collaborative nature of the research in many areas of physics and astronomy necessitate the importance of an efficient information system. The supply and the maintenance of such a system require up-to-date knowledge of scholars' information-seeking behaviour which is only achievable by researching this area.

### **1.3. Motivations for the study**

Although the researcher appreciates that there have been many studies in the area of information-seeking behaviour, there are still issues that necessitate further study:

a) *The need for deeper and closer understanding of subject differences in information-seeking behaviour.* Past studies have established subject differences among scholars in terms of their

information-seeking behaviour. We know that scholars in different fields have different patterns of information-seeking behaviour. The reliance on different information resources or different information seeking techniques varies among users from different fields and subjects. For instance it is well known that while monographs are important in humanities and social sciences, physical and life scientists rely more on journal papers (Brown, 1999a; Brockman et al., 2001). Studies by Nelson (2001), Rusch-Feja & Siebeky (1999), Smith (2003), Talja & Maula (2003), Tenopir and King (2002), Tenopir (2003), Tomney & Burton (1998) revealed differences in the use of electronic journals among scholars from different disciplines. However, the majority of the past studies tended to over-generalise their results by focusing on broad subject areas and attributing their findings to large academic or research fields. It is usual in the literature of information-seeking behaviour research to come across statements about information-seeking behaviour of, for instance, physical scientists, humanities scientists, social scientists, or if more specific, chemists or sociologists (see for example Tenopir and King, 2002). However, one must appreciate that these are broad academic or scholarly areas and each encompasses several more specific scholarly domains with their own research cultures and trends. Several authors have pointed out the need for narrowing the research focus. For example according to Bawden (2006:676) one of the changes that Wilson (1981) identified in his classic paper was a narrowing of research focus for in-depth studies of well-defined groups to determine the underlying factors of behaviour. Case (2002) believed that in studies of scholars' information practices, units of analysis should be even narrower than domains and specialties. Kling & McKim (2000) maintained that science should not be seen as a single type of activity: even in similar fields scientists can have very different research styles, communication patterns and information needs. Fry and Talja (2004) in their study on the use of e-journals showed that not only patterns of e-journal use vary across disciplines, but also within disciplines. These indicate a need for a narrower approach to the study of information-seeking behavior of the scholar, something that this study aims to pursue. Fry and Talja (2004) in their review of literature spotted this trend and highlighted it as one of the major limitations they identified in current approaches to the study of e-journals:

Comparative studies tend to be based on broad disciplinary groupings, such as the physical sciences, health sciences, social sciences, and humanities. Studies of this nature provide a broad picture of current usage patterns. However, they produce idiosyncratic results that do not adequately reflect epistemological activities within the knowledge producing communities that they attempt to represent.

An area such as physics consists of several narrower research areas including condensed matter, molecular physics, high energy physics, particle physics and chemical

physics among others. The level of interdisciplinarity, the scatter of the literature, the research process, the financial issues and industrial ties and many other factors might not be the same for all of these areas that are encircled under the umbrella of the term 'physics'. Each of these factors might have bearing on the ways scholars active in these fields communicate, conduct their research and seek information. As a matter of fact, there are indications in the research that some of these factors affect the information-seeking behaviour of scientists (e.g. Talja & Maula, 2003; for more details see Chapter 2). For example Bates (2002) has argued that domain size (the amount of topically relevant materials available relative to all materials in the area) and the degree of scatter in a domain are likely to influence search strategies in systematic ways. There is need for a better and deeper understanding of subject differences by means of intradisciplinary comparison and by taking a micro approach and looking at areas within a single discipline. Although there have been studies in the past that have focused on small fields, few studies have taken intradisciplinary approach to investigate similarities and dissimilarities within a broad discipline. Those that have chosen to study smaller areas normally compared subfields of different disciplines. For example Fry (2003) in her PhD thesis on scholarly communication compared high energy physics from physical sciences with social-cultural geography from social sciences and corpus based linguistics from applied sciences. To the best of the researcher's knowledge, this thesis is the first study to look at intradisciplinary similarities and dissimilarities within a single discipline.

b) *The need for a greater understanding of the information-seeking behaviour of the virtual scholar.* Scholarly communication is a live and evolving system, which has gone through intense changes during the past few years. These include the development of information technology with extensive applications in information services, evolution and growth of a variety of open access materials (institutional repositories, e-print archives), the provision of one-stop-shop sort of information services, and so forth. All these changes can affect information behaviour of the scholars. Today's scholars work in a rich digital information environment that virtually enables them to access most of their required information through a desktop computer. As Banwell and Gannon-Leary (2000) stated, the impact of information technology networks and electronic information systems and sources on academic users is potentially enormous, whether in support of research, teaching, publishing or communication.

Nowadays, the availability of digital information resources and services is phenomenal. Most scholarly journals and databases are now available online in electronic format. The digital technologies have improved the information services that academicians are provided with. This digitally provided information service requires a different kind of interaction by users compared to traditional print-based information resources and services. Digital

information services affect what Marchionini (1995) calls personal information infrastructures. Marchionini's notion of personal information infrastructures refers to an individual person's collection of abilities, experience, and resources to gather, use, and communicate information. He pointed out that information technology affect personal information infrastructures at all levels (1995:11-14).

Physics and astronomy are good examples of areas where virtual scholars work in a digital information-rich environment. There have been studies on physicists and astronomers' information-seeking behaviour, but they are either out of date (e.g. Ellis, Cox and Hall, 1993), have treated physics and astronomy as broad disciplines (Brown, 1999a) or have focused on a particular issue such as use of journal articles (Tenopir et al, 2005). There is need for a more comprehensive study of physics and astronomy as both are important fields of the basic sciences. This study, which fits into CIBER<sup>1</sup>'s Virtual Scholar programme<sup>2</sup>, tries to contribute in this area.

As far as the literature search by the researcher prior to the start of the project and during the course of the project has revealed, there have not been any studies with these characteristics - i.e. a recent, deep and holistic study - on physicists and astronomers' information-seeking behaviour at least during the last 15 years.

#### **1.4. Aims and objectives**

The primary aim of this research study is to investigate patterns of information-seeking behaviour of scholars in physics and astronomy. In particular, this study aims to derive behavioural patterns of information seeking activities of these scientists and to gain an understanding of what drives their behaviour. By undertaking the research at a micro-level (studying a group of scientists in one department of physics and astronomy) the study provides a detailed account of the scholars' information-seeking behaviour and activities. There has not been such a holistic and deep study on physicists and astronomers during the last 15 years (see Chapter 2), a period of significant changes.

The study compares the information-seeking behaviour of academics active in different research areas within physics and astronomy from an intradisciplinary perspective.

More specifically, the objectives of the study are:

---

<sup>1</sup> CIBER is a research group based in the School of Library, Archive and Information Studies at University College London.

<sup>2</sup> The Virtual Scholar is a research programme run by Centre for the Information Behaviour and Evaluation of Research (CIBER). It encompasses a range of studies on the supply and use of information in the academic community and the information seeking behaviour of academics. For more information visit: <http://www.ucl.ac.uk/slais/research/ciber/virtualscholar/>

- To investigate relationships between status (whether professors, students, researchers, and so on) and information-seeking behaviour. Understanding similarities and differences in information-seeking behaviour of users will help design better information systems. This will be particularly helpful for personalisation of information systems and services, something that currently attracts enormous attention from information service suppliers.
- To gain a deeper and up-to-date understanding of the information-seeking behaviour of researchers in different stages of a research process, and to investigate whether the techniques used for information seeking differ in different stages. This would clarify for example if scientists use different techniques when they are about to start a new research project compared to the techniques they apply during the project and also at the end of the project when they might be writing and publishing their results.
- To understand the reading behaviour of physicists and astronomers. The study examines the reading behaviour of physicists and astronomers in terms of the quantity of reading as well as some other aspects, such as the source of reading (whether articles or e-prints) and the role of abstracts. The study also investigates the relationship between academic status and research field, and reading behaviour and examines any relation between the amount of reading and methods used for identifying articles.
- To gain a deeper understanding of the publishing behaviour of physicists and astronomers. The study tries to find out more about the way physicists and astronomers go about their paper writing and publishing in order to understand the entailed information seeking activities in this process of writing and publishing. The interaction of physicists and astronomers with e-print archives is also investigated in this area.
- To identify difficulties they might face in their information seeking practices. The study will identify the barriers scientists face in their information seeking activities. The knowledge of the difficulties (if any) would help information professionals to enhance the scientific information system and improve information services for the scholar. Especially as the case study is an academic department at University College London, the findings will be helpful for the UCL library to improve its services for users at the department.

## 1.5. Research questions

Considering the aims and objectives of the study, the study has a set of research questions as follows:

- 1 Are there any important differences between different subfields of physics and astronomy with regard to different aspects of information-seeking behaviour, including methods used for keeping up-to-date and finding articles?
- 2 Are there any important differences in the information-seeking behaviour of physicists and astronomers according to academic status?
- 3 What are the techniques and methods applied to information seeking in the different stages of a research process in physics and astronomy?
- 4 What are the characteristics of the reading behaviour of physicists and astronomers?
- 5 What are the characteristics of publishing behaviour of physicists and astronomers?
- 6 What are the difficulties and problems that physicists and astronomers face in their information seeking activities?

## 1.6. Scope of the study

This study is confined to physics and astronomy and investigates the information-seeking behaviour of physicists and astronomers. The case study is the Department of Physics and Astronomy at the University College London. The department is one of the largest and oldest departments in its field in the United Kingdom. It is research oriented and highly ranked in the Research Assessment Exercise (ranked 5 in 2001 RAE). The department is involved in a wide range of research activities in different areas of physics and astronomy through four main research groups, each consisting of smaller groups and a number of associated research centres. The department is a research oriented one that at the time of data collection for this research (2005-06) had about 150 academic and research staff and more than 100 research students.

The reason why the researcher has studied both physics *and* astronomy is that these two fields are closely associated. Although there are areas in both physics and astronomy that they might not have mutual relations, there are also many areas in astronomy that are closely related to research in certain areas of physics and rely on it. The same holds true for certain areas in physics and their connection to research in astronomy. One may find it very difficult to draw clear boundaries between physics and astronomy, at least in the case of some subfields. The fact that these two fields together normally shape a single academic department, shows

their close ties. The other reason is that the case chosen for the study is an academic department of 'physics *and* astronomy'. It would not have been a wise decision if the research was restricted to part of the department. Cutting off half of the department from the study would have made it difficult to understand the research and information environment of the department that might have effects on the information seeking activities of its staff and students.

One must appreciate that the nature of research activities in academia is different from outside of academia. Hence, the information behaviour and scholarly communication of researchers in the academic environment and outside academia are different too. Tenopir and King (2002) mentioned several research differences inside and outside academic environments that affect scientists' communication and information activities. According to them:

The research performed in universities is usually basic and is largely funded by government, while applied research is largely funded and performed in industry. University scientists also tend to follow a single line (or related lines) of research throughout much of their careers, whereas in industry the focus of research changes more often depending on product requirement and managerial desires. Accordingly, university scientists tend to develop long-term, collegial relationships with other university scientists in their specialties and to communicate through invisible colleges both informally and formally. They also publish more and increasingly so, than the other scientists. Scientists in industry rely heavily on these publications, particularly in fields such as physical and life sciences (Tenopir and King, 2002:19).

The question that might be raised here is why the study focuses on an academic department and does not include researchers outside academia (*i.e.* industrial section). In the case of astronomers, they are less likely to work outside an academic environment. Astronomers tend to be academic or associated with academic institutions or at least work for governmental research agencies. This is due to the nature of the astronomy as a research field that tends to be an expensive field of research that requires extensive research facilities that are normally built by governments.

In the case of physics, although research in physics has a wide range of industrial applications and there are physicists that work in industrial sectors, they also tend to be academics. A study by Herschman (1977) found that physicists were more likely to be academic (rather than working for governments, not-for-profit organisations or industrial sectors) and more likely to be involved in research (rather than in other activities such as teaching, management and production). The situation still seems to be the same. According to the Bureau of Labour Statistics of the US Department of Labor, "most [of physicists and astronomers] work in areas in which universities, large research and development laboratories,

or observatories are located” (Bureau of Labor Statistics, 2006). Research by CIBER on the users of ScieceDirect physics journals also supports this. The users of ScienceDirect journals in physics appeared to be dominantly male academics (Nicholas, Huntington and Jamali, 2006).

Focusing the research on academics has some advantages for large and high prestige departments (such as UCL’s), which are expected to be centres of communication and all members of these departments, including those with little personal prestige, might benefit from this position. As Hagstrom (1970:108-109) pointed out, this is because most university scientists communicate more with their departmental colleagues than with the others, and they are often introduced to the work of scientists in other institutions by their departmental colleagues.

Unlike researchers in the industrial sector, academics are also involved in teaching and other activities. Studying a department also provides the opportunity to study the information-seeking behaviour of students who are future scientists. Hence this study focuses on research students along side faculty members. Undergraduate students are not included in this study.

The information-seeking behaviour of undergraduate students is expected to differ from faculty due to a number of factors. Undergraduate students’ information-seeking skills are not as well developed as faculty’s. They also have different information needs and seek information in a different context; they normally seek information to address the imposed questions by lecturers rather than self-selected questions (Rimmer et al., 2008). Graduate students, especially PhD students who are mainly involved in research activities, are more integrated in their departmental information environment and their communication and information-seeking behaviour is expected to be more similar to the faculty’s. The study of research (PhD) students’ information seeking-behaviour could shed some light on the way future scientists develop their information seeking skills.

The other point that needs clarification with regard to the scope of the study is the concept of information-seeking behaviour that is the focus of the study. Information-seeking behaviour is defined as “the purposive seeking for information as a consequence of a need to satisfy some goal” (Wilson, 2000:49). Although the focus of the study is information-seeking behaviour, the researcher touches on some elements of scholarly communication such as publishing papers as well as the issues of reading behaviour. This was because these elements provide contextual information for a better understanding of the information-seeking behaviour of scientists and help produce a clearer and more holistic picture of the subjects’ information-seeking behaviour.



### **1.7. Distinctive quality of the study**

This study has a few characteristics that make it somewhat distinctive among other research studies conducted on the information-seeking behaviour of physicists and astronomers. First of all, the fact that the case studied is one of the most prestigious departments of its kind in the UK and Europe with a wide range of research activities is an indication of the quality of the data used in this study. Studying such a large research-oriented department has advantages. Such departments have long well-established research cultures and tend to be centres of communication. The researcher himself belonged to the same university (UCL) that the studied department belongs to and this helped make the process of data collection proceed more smoothly and accurately. This was because obtaining the cooperation of the studied department was easier and the participants were also more inclined to cooperate with the researcher who belonged to the same university as they did.

As stated earlier, there has not been a holistic study on the information-seeking behaviour of physicists and astronomers for the last 15 years and this study tries to fill this gap.

Another distinctive quality of the study is that no other study has investigated information-seeking behaviour of physicists and astronomers (and to the best of the researcher's knowledge, of other scientists either) from an intradisciplinary perspective. The study adopts a micro approach which results in a clearer picture of the information-seeking behaviour of researchers in different subfields of physics and astronomy

The other distinctive quality of this work is its methodology. This mixed-methods study, as will be described in Chapter 3, encompasses four different data collection techniques including interview, survey, information-event card and desk research. It took advantage of both qualitative and quantitative approaches to overcome the shortcomings of each single data collection technique, triangulate the data and obtain a holistic and rich picture of the information-seeking behaviour of physicists and astronomers.

### **1.8. Thesis' outline**

The contents of this thesis are arranged in six chapters and twelve appendices. The current chapter sets the scene for the research and draws the foundations of the study. It provides the context and scope of the research and clarifies the aims and objectives of the research. The second chapter gives an overview of the literature. The reviewed literature covers information seeking studies as well as studies that are relevant to this research from methodological perspective or findings. The third chapter discusses the various methods applied for conducting the study, describes their specifications and justifies their use for this research. The

applied methods include interviewing, questionnaire survey and information-event cards. Furthermore the chapter discusses issues such as sampling, ethical issues and the limitations of the study. Chapter four illustrates the demographics of the research populations and the samples used for the different parts of the study. It presents the characteristics of the questionnaire survey sample, the interviewees and the participants in the Information-Even Card study. Chapter five is the results chapter. It presents the results obtained through the three main methods used in the study. The interpretation of the results, the conclusions and discussions on the findings of the study and their implications form the last chapter of the thesis which is the Discussions and Conclusions (Chapter 6). The thesis ends with the Bibliography and the Appendices.

### **1.9. Summary**

This chapter portrayed the outline of the study by explaining the aims and objectives, and rationale of the research and illustrating the scope of the study. In brief, the research is an attempt to derive behavioural patterns of the information-seeking behaviour of physicists and astronomers, who are good examples of virtual scholars working in and embracing a digital information environment (Gould & Pearce 1991, cited in Lawal, 2002; Wertman, 1999; Nicholas et al, 2005b). More specifically the study tries to better understand the information-seeking behaviour of physicists and astronomers with an intradisciplinary approach. The study compares subfields of physics and astronomy. Moreover, the study looks at the similarities and dissimilarities between people with different academic status in terms of their information-seeking behaviour. The information seeking activities of physicists and astronomers at different stages of a research project is another area that the research seeks to explore. Reading behaviour and publication patterns of physicists and astronomers and difficulties that physicists and astronomers face in their information-seeking behaviour are other issues investigated in this study. The focus of the study is mainly on research related information-seeking behaviour.

## CHAPTER 2 - LITERATURE REVIEW

### 2.1. Introduction

This chapter gives a review of the literature related to the scope of the current research, which is the information-seeking behaviour of scientists particularly physicists and astronomers. The literature in the area of information behaviour research is massive. Case (2007: 14) in his book estimated that the literature on information seeking extends over ten thousand publications in several distinct disciplines. The *Annual Review of Information Science and Technology* (ARIS&T) has also published several seminal review chapters on studies related to information behaviour areas since 1966 (Menzel, 1964; Herner & Herner, 1967; Paisley, 1968; Allen, 1969; Crane, 1971; Crawford, 1978; Dervin & Nilan, 1986; Hewins, 1990; Pettigrew, Fidel, and Bruce, 2001; King and Tenopir 2001; Wang, 2001; Case, 2005). Aiming to put the current research on the information-seeking behaviour of physicists and astronomers in context and making it possible to compare and relate the findings of this study to the past works, this chapter focuses on some specific studies that could be beneficial for understanding this study.

The chapter starts with a brief overview of the concept of information behaviour and information-seeking behaviour and then briefly discusses the approaches to information behaviour research and finally studies related to information-seeking behaviour of physicists and astronomers are reviewed.

### 2.2. Information behaviour

*Information behaviour* is a term used for a broad area that is dealt with in several disciplines such as library and information sciences and psychology. Referring to passive and active information behaviour, Wilson defined information behaviour as 'the totality of human behaviour in relation to sources and channels of information, including both active and passive information seeking, and information use' (2000:49). Using similar terminologies such as 'information seeking' and 'information need' Pettigrew, Fidel & Bruce in their classic review article defined information behaviour as 'the study of how people need, seek, give, and use information in different context including workplace and everyday living' (2001:44).

What can be inferred from all of these definitions is the broadness of the term 'information behaviour' and the fact that it encompasses some other information-related

concepts such as information need, information seeking and information use. Savolainen (2007) used the term ‘umbrella concept’<sup>3</sup> to refer to this broadness of the concept.

### 2.3. Information-seeking behaviour

One of the concepts under the umbrella of *information behaviour* is *Information-seeking behaviour*. Case (2002:5) considered it as a subcategory of information behaviour. Wilson (1999:263) in a seminal paper also proposed a nested model of information behaviour that shows how it encircles information-seeking behaviour (Figure 2.).

Several definitions have been presented for the concept of information-seeking behaviour. For example, to explain the concept of information-seeking behaviour Krikelas stated that it is:

...any activity of an individual that is undertaken to identify a message that satisfies a perceived need. In other words, information seeking begins when someone perceives that the current state of possessed knowledge is less than that needed to deal with some issue (or problem) (1983:6-7).

According to Marchionini (1995:5-7) information seeking is the process in which human purposefully engage in order to change their state of knowledge. It is a directed purposeful activity. Information seeking is a high level cognitive process. The common point between Marchionini’s and Krikelas’ definitions is their reference to the person’s state of knowledge and willingness to change it. Similar to Marchionini, another definition that highlighted the purposeful aspect of information seeking as an activity is Wilson’s. Wilson (2000:49) defined information seeking as ‘the purposive seeking for information as a consequence of a need to satisfy some goal.’ In more subjective terms, it can be considered as a process of locating resources or materials that fulfils information needs.

---

<sup>3</sup> Savolainen (2007) discusses another umbrella concept which is ‘information practice’. Information practice has been used as an alternative to information behaviour by some authors. Both information practice and information behaviour have been used to characterize the way that people generally deal with information. However, Savolainen (2007) discusses both of these two umbrella concepts and argues that they imply two different (though not yet most clearly defined) meanings. More about information practice will be said in ‘social approach’ section.

**Figure 2.1. A nested model of conceptual areas of information behaviour (Wilson, 1999:263).**

As we can see in both definitions by Wilson and Marchionini purposefulness is an essential attribute of information seeking as an action. Definitions and explanations by other researchers indicate similar attributes. Some researchers have tried to see it as a process and identify different stages and actions that it entails. Westbrook (1997) mentioned a group of five potentially interlocking actions that outlines the effort to seek information.

- 1      Needing: from the first hint that information may be of interest'
- 2      Starting: to work on the need
- 3      Working: on the need
- 4      Deciding: on the value of any results of working on the need
- 6      Closing: the effort to work on the need

In addition, he clarified that any of these actions may be the final one, or may be omitted, or may result in returning to an earlier action or in starting an entirely new effort. Although these actions are often sequential, these five actions can take place in virtually any order. However, Westbrook' account of information seeking is not a definition but rather a description of it as process or action.

## 2.4. Approaches to information behaviour research

Information behaviour is not expressed in isolation, but in the context of what people try to cope with in a specific time and in relation to one's information world (Solomon, 1997a, 1997b). There are several factors and elements involved in information behaviour including personal cognitive factors and social factors that have been investigated in the past research. Pettigrew, Fidel and Bruce (2001:46) classified and discussed research in information behaviour under three categories: 1) Cognitive approaches that cover those that examine the individual as the main driving force behind information behaviour; 2) Social approaches that examine frameworks that focus on the social context; and 3) Multifaceted approaches that cover those that consider multiple types of context, such as the cognitive, social, and organisational context. Allen (1996) also names a few perspectives on information behaviour research including cognitive, social, socio-cognitive and organisational. The domain analytic approach can be added to this list as well. Here the approaches of cognitive and social approaches as well as domain analytic approach that sound relevant to this research are discussed.

### 2.4.1. Cognitive approach

According to Cornelius (2002:406) the long development of the cognitive viewpoint in information science is generally believed to have started with De Mey (1977). The focus of cognitive approach is on identifying the cognitive characteristic involved and on the processing of information. De Mey maintains that the cognitive approach is

...that any *processing of information*, whether perceptual or symbolic, is *mediated* by a system of categories or concepts which, for the information-processing device, are a *model of his world* (1977:xvi-xvii).

Dervin and Nilan (1986) in their classic review article suggested that a new conceptual framework might be emerging within the literature of user studies, where the characteristics of the users rather than the systems, is in focus. Later reviews by authors such as Hewins (1990) and B. L. Allen (1991) confirmed this change. They both claimed that the literature on information need and use studies in the second half of 1980s was pervaded by cognitive research.

Slightly different descriptions and definitions have been suggested for the cognitive approach. For example Wilson (1984) stated that the cognitive approach:

...centres upon the idea of *meaning*. Meaning is involved not only in all aspects of information generation, transfer and use, but also in the way people define

themselves, their lives and their actions. The cognitive approach, therefore, draws attention to the need for a bridge between the meanings of everyday life and the information that may have relevance for everyday life. In this sense, of course, 'everyday life' is different for every person- for some it may involve research as an everyday activity, for others the practice of a profession, for others, involvement in business and commerce.

For Wilson, the central idea of the cognitive approach was the notion of 'human perception, cognition and structures of knowledge' (1984:197). Belkin (1990) explained this approach as being concerned with how an individual's state of knowledge interacts with what he receives, perceives or produces. Also, Ingwersen noted that the cognitive approach is concerned with explaining

... a *subjective* and profoundly *dynamic cognitive style* of information processing and cognition, ideally resulting in continuous changes of the models and the current state of knowledge for each device (1995:163).

The cognitive viewpoint, therefore, centres on identifying the characteristic features of a person that can explain variations in her information behaviour, where the main concern is the cognitive processing of information and a subsequent change of the mental image of the world and the knowledge structures. What this approach also suggests is that a person's information behaviour is a dynamic, constantly changing condition. Among the important theories that have been developed using cognitive approach, Dervin's (1983) sense-making model that turns attention to the primary cause of all users' activities, that is, cognitive discomfort, and Ingwersen's model (1984), which shows the relations among information and cognitive processes.

### 2.4.2. Social approach

The focus of social approach to information behaviour is on the impact of social life upon information seeking and evaluation of information or information sources. Pettigrew, Fiedel and Bruce (2001) in their review of the literature of information behaviour pointed out a shift in research in this area. They noticed that more studies started to emphasise the social aspects of information behaviour at the beginning of the 1990s.

Niedźwiedzka (2003) stated that researchers who apply the social approach see information users first of all as the members of a particular community, social category or group. They recognize the social placement or a professional role as the most important determinants of users' information behaviour.

Elfreda A. Chatman has been described as the originator of the shift from the cognitive approach to the social one (Pálsdóttir, 2005). Her studies have resulted in series of theories and concepts for studying information behaviour in the context of everyday life. While her

theories focused in the beginning on the issue of information poverty (Chatman, 1985, 1990, 1991a), they have since developed towards studying issues related to information behaviour in a more general sense.

She also developed the concept of small world (Chatman, 1991b) and the theory of life in the round (Chatman, 1999), which states that everyday life information behaviour is affected by the boundaries of the small world that people live in. According to the theory, the members will seek out only information that they believe is necessary in order to function within their small world, while information that is deemed as not necessary for their small world is ignored.

Later, she focused on normative behaviour, that is, social norms, worldview, social types and information (Chatman, 2000). Social norms are a set of standards that guide the members of the small world about patterns of behaviour and tell them what kinds of actions are expected of them, and what kind of behaviour is appropriate. Chatman further states that each person has a worldview that is shaped by the norms of the social world that he or she lives in. Worldview is a system of mutual beliefs that people belonging to the same small world have about the world around them and which affects their information behaviour. In dealing with their everyday life, people seek and use information according to the influences of their social environment. Their information behaviour is affected by what is believed to be normative behaviour in their small world. Among other models and theories of information behaviour that have been based on the social approach one can mention Katzer and Fletcher's (1992) model that shows specific information behaviour of managers.

The issue that is worth mentioning in the discussion on the social approach is the emerging term 'information practice' which is considered an umbrella concept similar to information behaviour. According to Talja (2005:123) information practice compared to information behaviour represents a more sociologically and contextually oriented line of research. Tuominen, Talja, and Savolainen (2005:328) pointed out that particularly from the constructionist perspective, the concept of information practice is preferred over information behaviour because it assumes that the processes of information seeking and use are constituted socially and dialogically, rather than based on the ideas and motives of individual actors. All human practices are social, and they originate from interactions between the members of community. In this way, the concept of practice shifts the focus away from the behaviour, action, motives, and skills of monological individuals. Instead, the main attention is directed to them as members of various groups and communities that constitute the context of their mundane activities. According to Savolainen (2007:119) the concept of information practice and its use has especially received more attention in the first half of the twenty-first century.



### 2.4.3. Domain-analytic approach

The domain-analytic approach is a paradigm in information science that was articulated by Hjørland and Albrechtsen (1995) and its development was a reaction to the cognitive approach that dominated the field at the time (Hjørland, 2007). It states that ‘the best way to understand information in IS is to study the knowledge-domains as thought or discourse communities, which are parts of society’s division of labor’ (Hjørland and Albrechtsen, 1995:400). In other words, domain analysis argues that it is more fruitful to view domains as basic units of analysis rather than focus on users in a generalised and context- independent manner (Talja, 2005).

To understand the domain analytic approach it might be helpful to consider the following points:

1) Domain analytic approach is not a totally new approach. The term ‘domain analysis’ was in use before the programmatic paper by Hjørland and Albrechtsen (1995). Researchers in computer sciences used the term ‘domain analysis’ even in 1980s (for instance Neighbors, 1980). The concept of domain analysis also can be seen in older studies. For example Saracevic’s (1975) subject knowledge and subject literature<sup>4</sup> view can be considered as one of the predecessors of domain analysis. In the field of information behaviour research any research that includes a sort of subject profiling is a kind of domain analysis. Hjørland and Albrechtsen (1995) in their paper on domain analysis listed some ideas by other scholars as contemporary approaches of a related nature. Among them they have mentioned the opinion by Patrick Wilson (1993) indicating that unit of study in information science should be the speciality, not the individual; or Saracevic’s (1975) concept of subject knowledge view of relevance, or Thomas J. Froehlich’s (1994) suggestion implying that the foundations of information science must be in social epistemology. Talja (2005) also mentioned a few works such as Diana Crane’s work (1972) on invisible colleges, T. J. Allen’s (1977) book on dissemination of technical information, and William Garvey’s (1979) book on scholarly communication as classics of domain analytic approach.

However, in information science Hjørland (2002b:259) claimed ‘to be the first person to have used this term and the underlying theory and methodology.’ He used the term first in Hjørland (1993). Also it was not before Hjørland and Albrechtsen’s (1995) paper that

---

<sup>4</sup> Subject knowledge view of relevance considers the relation between the knowledge on or about the subject and the subject. The subject literature view of relevance is closely related to the subject knowledge view and considers the literature, or the relation between the literature and a topic (question) on the subject (Saracevic, 1975).

researchers started to use this approach in the sense described by the authors for research in information science. For example Beghtol (1995) did a 'domain analysis' study on the literary warrant of bibliographic description of fiction studies which was published before Hjørland and Albrechtsen (1995), however her approach was not exactly what was explained in Hjørland and Albrechtsen's. What Hjørland's articulation of domain analysis adds to it is the attempt at finding an explanation of peoples' use of information. How information is selected, what relevance criteria are at play etc. For example Hjørland' (2002b) showed that scientists' relevance criteria are closely connected to their epistemological views.

2) The articulation of domain analytic approach by Hjørland was a reaction to the cognitive view, which focuses on individual minds rather than social groups (Hjørland, 2007).

#### *What is a domain?*

It needs to be said that as domain analytic approach is still in its infancy and is being developed, there seems to be no firm consensus on the definition of domain. Hjørland (2004) explained that a domain 'may be a scientific discipline or a scholarly field. It may also be a discourse community connected to a political party, a religion, a trade or a hobby'. For example, chemistry can be considered as a domain which is both defined ontologically as about entities in the world and socially as the group of people studying that field. Normally academic disciplines are considered as domain, however researchers develop operational definitions for their research and a domain can be considered as a specific research area within a broader academic discipline. For example, Fry considered intellectual fields as the units of analysis in her domain analysis study (Fry, 2006).

Tennis (2003) with a different approach sees the definition of domain as an open question and a matter of operationalization. Therefore he proposes two axes that can be used by a domain analyst in operationalizing his or her definition of a domain. The two axes are:

Areas of modulation that sets parameters on the names and the extension of the domain. The extension of the domain is its total scope. Area of modulation must state a) the totality of what is covered in the domain analysis - the extension and b) what it is called – its name.

Degrees of specialization that qualify and set the intension of a domain. The degree of specialization must a) qualify the domain – state its focus and b) state where the domain is positioned against other domains – its intersection.

However, Hjørland (2007) found Tennis's approach to the definition of domain problematic. Tennis (2003) maintains that 'the notion of domain must be defined in a transferable definition – one that can be used by more than one researcher, to allow for a shared understanding of what the object of domain analysis is.' Unlike Tennis, Hjørland and

Hartel (2003) argue that ‘while it might be easy to select one turn-key definition of a domain, such a definition will always be more related to one view or paradigm, and relatively unsatisfactory for other paradigms.’

In conclusion, the literature indicates that the concept of domain is still somewhat open to interpretation and further development, and researchers operationalize its definition. Although the popularity of this approach as a conceptual framework for studies in the field of information behavior has been on increase during the last few years and more studies have been conducted using this approach (cf. Brown, 1999a; Brockman et al., 2001; Fry, 2006; Fry & Talja, 2004; Tenopir et al., 2005), still some researchers (for example Palmer, 1999; Bates, 2002; Hjørland, 2002b; Fry and Talja, 2004) believe that this approach is in its infancy. One of the Hjørland’s (2002a) eleven approaches to domain analysis is empirical user studies. However, the explanation presented for this approach is broad and general and unlike the other approaches for which he presented methodologies, no clear methodologies were presented for user studies.

## **2.5. Models of information-seeking behaviour**

Since the emergence of the user-centred paradigm in the information behaviour literature which was observed by Dervin and Nilan (1986), there have been consistent efforts for conceptual and theoretical enrichment of the field. By adopting the qualitative approach to research in information behaviour and borrowing theories from the social sciences, researchers have been able to develop theories and conceptual frameworks. Theories have predictive value and conceptual frameworks may be used for the examination of definitions and assumptions.

Using these approaches, researchers have developed a few models of information behaviour and some models of information-seeking behaviour in order to better understand the behaviour and activities of users. A model, according to Wilson (1999:250), is a framework for thinking about a problem and may evolve into a statement of the relationships among theoretical propositions. Wilson (1999:250) differentiated between models of information behaviour and models of information-seeking behaviour while in some other publications (for instance, Case, 2007:120-138) there is not such a differentiation. According to Wilson (1999: 251) the number of general models of information behaviour appears to be few. One of them is Wilson’s (1981) model, of which a modified variation was presented in Wilson (1999:251). See Figure 2.2. Information need plays a central role in Wilson’s model.

**Figure 2.2. Wilson's model of information behaviour (Wilson, 1999:251).**

Models of information-seeking behaviour are more numerous. Here we discuss Ellis's model, a well-known model that has been claimed to be applicable to all disciplines (Ellis, Cox & Hall, 1993:359) and has been the basis of some later studies. This model was chosen for it is applicable to the use of different information resources and it has already been studied on physical scientists (who are the subject of the current study), although more than a decade ago (Ellis, Cox and Hall, 1993). There are some other models but they are restricted to a particular group of information users and not pertinent to the population of this study. For example the model by Leckie, Pettigrew and Sylvian (1996) is restricted just to professionals (engineers, lawyers, and health care professionals) or the one by Baldwin & Rice (1997) just deals with the small community of security analysts. Kuhlthau's (1991) model which is called 'information search process' is mainly about searching and is not suitable for the broader activities of information-seeking behaviour. Marchionini's (1995) and Ingwersen's (1996) models are restricted to a specific task, which is searching electronic information. The stages such as 'choosing a search system' and 'query formulation' in Marchionini's model make it more of a searching behaviour model.

### 2.5.1. Ellis's model of information seeking strategies

Ellis (1987) used in-depth semi-structured interview technique and adopted grounded theory to conduct a qualitative study on the information-seeking behaviour of social scientists. His primary goal was to apply the behavioural approach for designing an information retrieval system. His study resulted in a model of information seeking behaviour of social scientists. The model had six features:

- Starting: the means employed by the user to begin seeking information, for example, asking some knowledgeable colleague or identifying a key paper to commence the search;
- Chaining: following footnotes and citations in known material or “forward” chaining from known items through citation indexes;
- Browsing: semi-directed or semi-structured searching in an area of potential interest;
- Differentiating: employing differences in the nature of the source material to filter material;
- Monitoring: maintaining awareness of development in a field through regularly following particular resources;
- Extracting: selectively identifying relevant material in an information source;

Ellis later expanded this model by testing it on physical scientists (physicists and chemists) (Ellis, Cox & Hall, 1993) and Engineers (Ellis & Haugan, 1997). The study of physical scientists resulted in the addition of two more features to his model. The two are:

- Verifying: activities associated with checking the accuracy of information;
- Ending: activities characteristics of information seeking at the end of a topic or project, for example, during the preparation of papers for publications, or it can be defined as “tying up loose ends” through a final search.

Ellis believed that this set of common characteristics of information-seeking behaviour of researchers is applicable to all disciplines. Moreover, Ellis pointed out that this model does not define the interactions and interrelationships between the categories or the order in which they are carried out (Ellis, Cox & Hall, 1993:359). He argued that the interrelation or interaction of the features in any individual information seeking pattern depends on the unique circumstances of the information seeking activities of the person concerned at that particular point in time. He added that the relationships between the features of the model can only be indicated in the most abstract and general terms unless there is reference to a particular

information seeking pattern (Ellis, 1989:178). However, Wilson (1999) suggested that the features in Ellis's model could be presented in a sequential way. He presented the following diagram (Figure 2.3).

**Figure 2.3: A stage process version of Ellis's model (Wilson, 1995: 255).**

Ellis's model later was revised by Meho and Tibbo (2003). They used semi-structured email interviews to investigate the information-seeking behaviour of scholars active in the field of stateless nations (a research area in social sciences) such as Kurds. Their sample was international and consisted of about 65 academics. Their research resulted in the addition of four new features to the model: accessing, verifying, networking, and information managing.

- The issue of access was regularly brought up by participants in Meho's and Tibbo's (2003:581&583) research, because a great deal of information was identified through bibliographic databases, personal contacts, publisher's catalogues or backward and forward chaining. Maybe the topic of stateless nations was particularly vulnerable in this sense and caused the problems with access to researchers.
- Verifying is characterized by activities associated with checking the accuracy of the information found. The study participants emphasized these activities primarily because of the political and sensitive nature of their research topics.
- Networking is characterized by activities associated with communicating, and maintaining a close relationship, with a broad range of people such as friends, colleagues, and intellectuals working on similar topics, members of ethnic organizations, government officials, and booksellers. Perhaps participants of the study created, or participated in, networks not only to build collections or gather information, but also to share information with members of these networks

- The need and importance of filing, archiving, and organizing the information collected or used was mentioned by several participants. Knowledge is not always immediately obtained or applied. It needs to be gathered, digested, organized, and stored for future use. According to the study participants, personal collections not only provide them with easily accessible materials, but also with materials that are organized or classified in a way they understand (Meho and Tibbo, 2003).

As a result of the new features identified in the study, Meho and Tibbo (2003) developed a new model, which, unlike Ellis's, groups all the features into four interrelated stages: searching, accessing, processing, and ending.

**Figure 2.4. Stages in the information-seeking behavior of academic social scientists (Meho and Tibbo, 2003:584).**

As Meho and Tibbo (2003:584) explained (see Figure 2.4) a number of activities could occur at each of the first three stages (searching, accessing, and processing). During the search stage, researchers might use starting, chaining, browsing, monitoring, differentiating, extracting, and networking activities. During the accessing stage, researchers could make a decision whether to proceed to the processing stage or return to the searching stage. This decision was likely to be based on the success or failure in obtaining needed material and gaining access to various information sources. During the processing stage, researchers might use chaining, extracting, differentiating, verifying, information management, synthesizing, analyzing, and writing activities. These last three activities (synthesizing, analyzing, and writing activities) did not belong to actual information-seeking behaviour, but they were mentioned here obviously because they were so essential to a researcher's work.

Meho's and Tibbo's information seeking model, that is a kind of research cycle, has interested some researchers (Poteri, 2007). Its cyclical and dynamic nature has been

appreciated. However, more empirical research has been suggested to confirm the stages described in the model. Lindström (2005:6, cited in Poteri, 2007:29) found the issue of access problematic because it is not clear whether it included only accessing formal sources through different facilities or technologies. Social institutions and communities might build up barriers to access, or support access as well. Therefore he argued that Meho and Tibbo had a rather traditional and technology oriented view on access. He also paid attention to how small a role contextual factors seemed to play in Meho's and Tibbo's model. According to Lindström (2005:10), Meho and Tibbo did not really examine how much researchers' positions and skills affect information seeking.

The latest development in Ellis's model is the study by Makri, Blandford and Cox (2008). They used Ellis's model as a lens to analyse and make design suggestions based on the information-seeking behaviour of 27 academic lawyers, who were asked to think aloud whilst using electronic legal resources to find information for their work. They identified similar information-seeking behaviours to those originally found by Ellis and his colleagues in scientific domains, along with several that had not been identified in previous studies, such as 'updating'. They also presented a refinement of Ellis's model based on the identification of several levels that the behaviours were found to operate at and the identification of sets of mutually exclusive subtypes of behaviours. Their study illustrated that Ellis's model is useful for informing design. While Ingwersen and Järvelin (2005) had already asserted that Ellis's model is not suitable for providing design insights because Ellis's characteristics provide types of activities that users might want to accomplish through the systems and not any *direct* design specifications for interactive systems.

They maintained that their study validated Ellis's model through a new research method of Contextual Inquiry (which includes a naturalistic observational element as well as an in-depth interview element). According to them this is particularly useful because all of the previous studies that identified information-seeking behaviours (i.e. those by Meho & Tibbo, 2003 and Ellis and colleagues) only used a research method based on semi-structured interviews. This means that previous studies have only been based on participants' *reports* of the behaviour that they display as opposed to *observed* behaviour. They identified multiple levels at which the lower-level behavioural characteristics can operate with regard to electronic resources – at the resource level (i.e. at the level of the electronic resource itself), the source level (i.e. at the level of an information source or sources within a particular electronic resource), the document level (i.e. at the level of a document or documents within a particular information source), the content level (i.e. at the level of content within a particular document) and the search query/result level. Selecting, updating, recording, and collating &



editing were the types of behaviours that the authors claimed that the previous studies had not identified.

Case (2007:122) maintained that Ellis's model makes "no claim to consider many of the factors and variables generally considered in information seeking research: the type of need and what sort of information or other 'help' might satisfy it; or the availability of sources and their characteristics".

## **2.6. Information-seeking behaviour in academic environment**

As Liull stated the academic environment is 'the environment which emphasises the learning or discovery mode motivated by individual's commitment to expand the human knowledge base' (1991:84). According to Marchionini (1995:32) academic environment consists of administrative staff, students (undergraduates, postgraduates and research), faculty members (professors, readers, lecturers), and research staff (postdoctoral researchers and research fellows).

Academic environment can be considered as a context of information-seeking behaviour. According to Poteri (2007:12) the concept of context is often used in information studies, but seldom defined. Sonnenwald (1999:178) suggests that context is "the circumstances in which a particular event or situation occurs". In her view, examples of contexts include working life, family, citizenship, university, or school. Each of these contexts has boundaries and privileges as gained by participants. There must be some shared understanding about the context. Defining different contexts is a complicated task because contexts are not discrete. For example, a faculty member may also be a teacher and an administrator. An outsider cannot easily determine when a faculty member is acting in the context of teaching, or in the context of administration (Sonnenwald, 1999:178).

The main activities in an academic environment are teaching and conducting research. Each of these activities involves a number of more specific tasks and each task can be a source of information need. Information seeking depends on interaction among several factors including information seeker, task, system, domain, setting, and outcome. Moreover, in an academic environment, which concerns the researcher in this research, each of these factors has a limited range of possibilities. For example, tasks of information seekers in academic environments are presumably limited to research, teaching, writing articles, and so on. The context of information seeking in academic environments has its own characteristics. Context of information seeking can be described by means of many different parameters such as the time and place of appearance of information need, the time for information seeking, types of

participants of the seeking process, for example, their demographic, social, professional, educational and behavioural characteristics, the purpose of information seeking, the concrete task for which this information is looked for, the processes and situations of information seeking, and many others (Gaslikova:1999).

It should also be born in mind that there are differences in the information seeking habits and behaviours of academic scientists and scientists active in industrial sectors. The difference has been discussed for long time (Herner, 1954; Tenopir and King, 2002). Herner (1954) many years ago stated that researchers performing academic duties make greater use of formal information channels or sources, depend mainly on the library for their published material, and maintain a significant number of contacts outside of the organisation. While researchers performing industry duties make greater use of informal channels or sources, depend on their personal collections of information and colleagues for information, make significantly less use of the library than do their counterparts, and maintain fewer contacts outside of the organisation. Some of these differences seem to still exist in the 21<sup>st</sup> century as Tenopir and King (2002) also described some differences in the research performed in universities and outside academic environments and they made similar points. According to them while the academic research is usually basic and is largely funded by government, the applied research outside academia is largely funded and performed in industry. University scientists tend to develop long-term, collegial relationships with other university scientists in their specialties and to communicate through invisible colleges both informally and formally. They also publish more and increasingly so, than the other scientists. Scientists in industry rely heavily on these publications, particularly in fields such as physical and life sciences.

## **2.7. Physicists and astronomers**

There is a considerable body of literature on scientific communities from the perspective of the sociology of science. A good example is the work by Becher and Trowler (2001) on academic disciplines. Although it is appreciated that studying scientific communities from that perspective helps understanding scholars' communication and information behaviour, this area is out of the scope of this study and therefore it is not dealt with here.

The existing literature in library and information science on physicists and astronomers is not evenly distributed over different aspects of their scholarly communication and information-seeking behaviour, the main focus is on their scholarly communication. While for example the literature on the preprint culture of physicists and astronomers is rich, their information-seeking behaviour has not been well-covered. The following sections review the literature on different aspects of information-seeking behaviour and scholarly communication of physicists and astronomers.

### 2.7.1. Information-seeking behaviour

Not many studies have been done on the information-seeking behaviour of physicists and astronomers and the methods they utilise for finding information. A rather old study by Ellis, Cox and Hall (1993) before the popularity of the Web-based information services investigated the information-seeking patterns of a group of social scientists, physicists and chemists using the grounded theory approach. They did not find fundamental difference among these groups, surprising perhaps. Therefore they concluded that overall differences between the information seeking activities of the chemists, physicists, and social scientists seemed more a difference of emphasis than of a fundamental difference in behaviour. They identified five main features for the information-seeking behaviour of these groups with slightly different terminologies. For the physicists these five core features were: initial familiarization, chasing, source prioritization, maintaining awareness and locating. These five features are known as Ellis' model of information-seeking behaviour (for more information on Ellis' model see section 2.5.1)

Regarding applied methods for finding less recent information, a survey of astronomers, chemists, mathematicians, and physicists at the University of Oklahoma by Brown (1999a) found that physicists and astronomers used citations at the end of articles (94%), retrospective searching of indexing/abstracting tools (56%), personal communication (50%) and browsing older volumes (19%). Eighty one percent of respondents in the field of physics/ astronomy said that they photocopied library's copy for obtaining journal articles, 75% read library's copy, 44% used free electronic copy, the same percent used interlibrary loan, 38% had personal subscription, and 19% used library's electronic subscriptions.

Nicholas et al's (2005b) surveyed the users of the Institute of Physics (IoP) journals to report on the views and attitudes of physicists around the world in relation to what they get and want from the journals system at a time of change and uncertainty i.e. at a time when new publishing models, like open access journals, are being proposed. Their aim was in particular, to assess how they use electronic products and services. The findings showed that the most frequent method used to locate journal articles was visiting a journal's website. Respondents said they were most dependent on visiting a journal's web site for finding articles followed by the library. Younger respondents were more likely to rely on the Web of Science, while older respondents were likely to depend on their personal collection. In general, younger users were more likely to depend on and use online methods. The most important web site proved to be the arXiv e-print server and this was followed by Elsevier Science Direct.

Another study by Nicholas, Huntington and Jamali (2006) that utilised both log analysis technique and online survey to study the information-seeking behaviour of ScienceDirect's authors (as users) revealed some information about physicists in comparison to scientists in other fields.

- Physicists, compared to respondents from other subjects, were more likely to be browsers than searchers - they obtained much information by requesting journal homepages and journal issues.
- Compared to the respondents from other subjects, Physicists were more active, making a higher number of requests in a session.
- Physicists favoured PDF over the HTML format and they requested more abstracts than users in the other subjects.
- Women were much less likely to search the Physics literature than men and the proportion was much higher than for the other subject fields 8% as compared to 25%.
- In the case of Physics less use came from the university sector and more came from research organisations, but in either case the differences were not dramatic.

In terms of the percentage distribution of item requests by subject in physics, articles accounted for 31% of Physics item requests. This was slightly lower than the 33% article requests recorded for other subjects. Visits to the journal homepage were higher in Physics as compared to the other subjects, but not significantly (18 and 14% respectively). The numbers of requests for journal homepages, journal issues in Physics were also greater. On the other hand, Physicists made fewer requests for the search menu. These results show that users from Physics were browsers. Physics recorded one the highest frequencies of sessions requesting only PDFs (43%).

The researchers also lined attitudinal data obtained through survey to log data and demonstrated that:

- Respondents viewing Physics journals were less likely to agree with the statement 'the quality of an article is determined by the journal' compared to other subjects. Physics scored 2.82, the fourth lowest score. Economics (3.43) Environmental Science (3.63) and Material Science 3.27 scored highly by this question, while Mathematics (2.57), Social Science (2.73) and Engineering (2.85) recorded low scores.
- Those people using Physics journals were less likely to agree with the statement 'it is more important to publish in a prestigious general journal, than a MORE

appropriate specialised journal'; their score was 2.65 while those accessing Economics (3.65), Mathematics (3.44) journals tended to agree with the statement.

ScienceDirect users had a range of search options. Compared to the other users those from Physics were more likely to use the all full-text resources option, 55% did so. Thirty eight per cent of users from Physics employed the search within journal. Physicists recorded the highest rate of use of ScienceDirect Volume Issue Alerts (html); 26% of those entering the site via an external link used it.

A few of the past studies investigated the issues that could affect the searching methods used for identifying journal articles. Bates (2002) for example hypothesised that scholars in high scatter fields use chaining and browsing as their primary search methods, whereas directed keyword searching is a more effective method for finding relevant materials in low scatter fields. A study on the use of electronic journals by Talja and Maula (2003) showed that A high scatter lead to a more intensive use of both journal and reference databases. Vakkari and Talja (2005) surveyed users of FinELib. The findings showed that scatter has a significant influence on the use of electronic information sources. Researchers using literature mainly from several fields used more databases of all types compared to colleagues who used literature mainly from their own field. The increase in scatter increased only the importance of searching in reference databases. The number of relevant databases is higher in high scatter fields implying greater effort in keeping up-to-date. In some fields, especially those with a higher degree of vocabulary control, directed searching across fields is greatly facilitated by mixed-journal databases containing journals from several fields. In other disciplines, researchers in high scatter fields probably reduce their search load by first searching databases for references and then continue to the full-text journals.

However their subject categories were very broad, for example physics, maths, forestry, chemistry, food industry, home economics and some other disciplines were are grouped together as Natural Sciences.

### 2.7.2. Reading behaviour

The studies on the reading behaviour have covered three aspects including the amount of reading, the sources of reading and the reading process itself. In terms of the amount of reading, a survey study in U.S.A. in 1981 showed that physicists read about 20,000 separate copies of articles from 19 American Institute of Physics journals; 4,500 of them were preprints. Physical sciences authors averaged distributing 110 preprints per article (King, McDonald & Roderer, 1981). More recent studies by Tenopir and King (2002) found that the number of readings per scientists across all work fields increased from an average of 100 articles per year

in mid-1990s to 130 per year in early 2000s. A survey in 2000 showed that physicists read an average of 204 articles per year and spend 153 hours per year reading on average.

With regard to the sources of reading and their importance, a study in early 1980s showed that in terms of the importance of sources of reading, current journal literature was of high importance for astronomers and physicists, but they also depended on preprints of articles that may finally appear in a journal (King and Roderer, 1982). Brown (1999a) surveyed astronomers, chemists, mathematicians, and physicists at the University of Oklahoma and found that physicists and astronomers mostly used current journals and had reliance on preprint archives. Mathematicians and physicists / astronomers used personal communication and conference attendance but chemists used current awareness services. This study also shed some light on the different sources of reading used for different purposes. It turned out that for teaching purposes, physicists and astronomers relied 93% on Textbooks, 40% on journals, 6% on preprints, 6% on conference attendance and 13% on personal communications, while these figures for research purposes were respectively 33, 87, 67, 60 and 33 percent.

A year after Brown's study, Cho (2000) investigated the sources of reading by physicists and concluded that astrophysicists have virtually replaced journal reading with regular inspection the Astrophysics archive at arXiv.org. This was not surprising. In Fact, Taubes (1993) and Ginsparg (1994) argued physicists studying high energy particle theory were among the first scientist to make use of electronic information delivery when the Los Alamos National Laboratory electronic preprint archive was established in 1991. However, Brown (2001b) discussed that although physicists and astronomers rely on e-prints for information, the use of printed articles in the top tier physics and astronomy journals was not diminished during the time following arXiv.org's establishment. Her study indicated that both the total numbers of yearly citations to and the ISI (Institute of Scientific Information) impact factors of the 37 journals studied had remained fairly constant, or had slightly risen, since the beginning of arXiv.org in 1991 until 1998.

Tenopir and colleagues have pioneered studies on the reading behaviour of scientists since 1977. They have conducted a series of surveys on scientists (mainly academics) over almost three decades that their findings give a good longitudinal picture of changes in the pattern and amount of reading done by scientists. Their studies on the readership patterns showed that they are enthusiastic users of electronic journals (Tenopir, 2003). Tenopir et al. (2005) investigated the use of electronic journals by astronomers of the American Astronomical Society and compared the results with the results of the surveys of other scientists. Although the results were published in 2005, the two surveys were conducted in 2001. They conducted two surveys with a random sample 2200 astronomers each and received

respectively about 23 and 22 percent response rate. The findings of the surveys indicated that the availability of a mature electronic journals system from their primary professional society has surely influenced their early adoption of e-journals. This influence has different aspects as nearly four-fifth of their readings is from electronic sources, compared with just over one-third of readings by other scientists. Also electronic articles tend to be read more thoroughly than the print versions; and this is reflected in the average time spent reading (35 minutes per reading vs. 27 minutes). The authors maintain that this difference is an indication that the information content is at least as valuable as the print versions for astronomers. The other possible effect of the access to electronic reprints is that they rely much more on the citation chasing for finding articles. A difference between Tenopir and colleagues' study with the others' is that Tenopir and colleagues' study shed some light on the age of the material read. It showed that astronomers have more re-reading than other scientists, which may be due to the ease with which they can retrieve old articles electronically. About 63% of the articles they read are one year old. Compared to other scientists astronomers read more for primary research and less for current awareness. They are observed to identify fewer of the articles they need through browsing than other scientists do. For primary research they rely more on online searching while for background studies and current awareness their reliance on browsing is more than on online searching and other means.

Little research has been done on the process of reading itself. The only study on physicists in this area is the old study by Bazerman (1985, 1988). He interviewed seven research physicists to understand the process of reading and to find out about their reading behaviour and the way they decided what material to read. His research found that in making the early choices of what articles to read, physicists call on personally organized knowledge. This extends beyond textbook knowledge of accepted facts and theories to include dynamic knowledge about the discipline's current practices and projections of its future development. The knowledge even includes judgments about the work of colleagues. Bazerman explained that:

In selecting the range of reading the physicists must, of course, have a sense of the various fields of current work. Moreover, in deciding the urgency of reading the physicists must rely on an image of how rapidly work moves in their fields. (1988:239)

He found out for instance that all the pure theoreticians and experimental biophysicists went to the library at least once a week to search the tables of contents of newly arrived journals because they perceived their fields as moving rapidly and they must keep current to do adequate work. Both physicists in remote sensing, however, chose less timely methods of search—one using Current Contents and the other using abstract indexes. When he questioned

the interviewees about the slowness of their search techniques, both remote sensing physicists said that their field did not move fast enough for that to matter.

The scanning processes of these physicists give evidence about how deeply these schemas are impressed in the subconscious, the subjects scan so rapidly over tables of contents that they cannot give conscious thought to each title. Rather, certain words seem to trigger the attention and make the scanner question a particular title more actively (Bazerman, 1985, 1988).

### 2.7.3. Journals, e-prints and open access

Physicists and astronomers have played a significant role in the scholarly communication and publishing, especially in areas such as e-print culture and electronic publishing. They are renowned for having one of the, apparently, most efficient information systems (Nicholas et al, 2005b) and the best organised literature in sciences (Gould & Pearce 1991, cited in Lawal, 2002). They are known as innovators in methods of scholarly communication (Wertman, 1999). In fact these two fields were among the leaders in the experimental publishing models and scholarly communications, such as pre-print and e-print archives and open access publishing were started and embraced. For example Institute of Physics Publication, as one of the pioneers in electronic scientific publishing, started making its journals available online as early as September 1994 by launching Classical and Quantum Gravity on three types of servers including List, Gopher and World Wide Web (Singleton, 1997:152) and it was the first major publisher that made all its thirty three journals available online in January 1996 (Dixon, 1999:3).

Traditionally like most of other scientific areas the literature of physics and astronomy is largely based on journal publications (Gould and Pearce, 1991 as cited in Lawal, 2002). However, physicists and astronomers have a long and rich preprint culture. The practice of sending out preprints, although common among many fields of science, has long been established among physicists and many scholars have remarked about this distinctive preprint culture in physics (King and Roderer, 1982; Hurd, 1996). Physicists have used preprints for over thirty years (Brown, 2001a & b) and the initial electronic equivalents of preprints started about 1991 on a very small scale but rapidly grew to a current astronomical level of tens of thousands transactions per day (Valauskas, 1997). Physicists and astronomers are heavy users of e-print archives (Kling & McKim, 2000; Fry, 2003). The statistics of arXiv.org monthly submission (arXiv, 2007) shows steady increase in the submission rate since 1991.

Lawal (2002) explained why physicists have high usage of e-print archives. She stated that:



Theoretician physicists depend on the work of their predecessors. The information most important to them is often too recent to have been published; hence they use e-print archives. Experimentalists are more concerned with the way in which experimental procedures are carried out. Experiments in high-energy physics are very expensive; often physicists cannot wait for formal publications. High-energy physicists have depended on preprints for a long time... Preprints are most valued in physics because they provide an instantaneous publication channel. Physics is also collaborative in nature. It is not unusual to find a physics paper with over one hundred authors. These reasons, with a long existing e-print archive explain why physicists have the highest use.

The only qualitative study on the use of e-print archives by physicists in the past is the study by Wertman (1999). She interviewed twelve physicist and chemists at the University of Maryland. Besides cross disciplinary differences between physics and chemistry, she found differences within physics. The study concluded that particle physicists and condensed matter physicists used e-print archives more than the other subfields of physics did. It was also demonstrated that theoreticians, especially, gravitated to the archives because the medium supports this wonderful way of communicating ideas quickly and fluidly, almost replicating a conversation. Experimentalists were less likely to use the archives, and it may be because they do not want to wade through unrefereed papers. It was thought that one reason for more use of e-print archive by theoreticians might be because recognizing poor science may be easier in the theoretical fields that spawned the first bulletin boards than in experimental fields, where a reviewer has to evaluate experimental design and statistics as well as mathematical reasoning.

Wertman's (1999) study also showed that while graduate students and post-doctoral researchers were called on to navigate the technical difficulties of electronically submitting papers, it was the older more established scientists who promote usage.

Journal publishing has been important in physics and publishers active in the field of physics have paid attention to physics community as the readers of the journals. Institute of Physics Publishing (IoPP) is one of the main publishers in the field of physics that have investigated different aspects of scholarly communication of physics community. IoPP conducted a world wide survey of physicists in mid 1990s but the results were not published. However, Singleton (1997) mentioned some of its results in a presentation. They sent a large questionnaire to over 13,000 physicists around the world. Some 3500 were completed and returned. The main focus of the survey was issues concerning scholarly communication and publishing. The participants were asked all sorts of things in their roles as authors, referees, readers, influencers, and purchasers. Physicists were asked about their knowledge of e-print servers which at the time already existed in most major subfields of physics. Just over half (54%) did not know whether there was one in their field, but over three fourth of these (and those who thought they did not have one) said they would like one. They were asked about what should happen to an e-print when a final version was published. Forty four percent

thought it should be deleted immediately and the 56% disagreed. Most of those who disagreed asked for later deletion. As to electronic journals, the interesting finding was that they mentioned full peer review as one of the top five important features they wanted to see in e-journals. The other four features were searchability of abstracts, possibility to printout locally, possibility to browse table of contents, and typeset quality maths. They also were asked about the effect of refereeing on the papers. Just 8% said that their papers had not improved as the results of refereeing and the rest maintained that their papers had improved with 15% of them said the papers improved between 79-100 percent. As referees, 14% said that the material always improved as the result of their refereeing. As the IoPP survey showed the popularity of technologies such as email among physicists, two other studies (Walsh & Bayma, 1996; Barry, 1997) almost at the same time as IoPP survey showed that use of bulletin boards, distribution lists and electronic mail was common in physics.

More recently another unpublished survey was conducted for IoPP. Physicists who had published in the Institute of Physics' (IoP) journals were surveyed by Nicholas et al (2005b). The study showed that the most well known journal was *Physical Review Letters*. In terms of importance as a research tool, the *Journal of Physics* series did not perform very well and were outperformed by both *Physical Review* and *Physica* series. In terms of where respondents published, the three journals recording the highest percentages were the *Physical Review B*, *Physical Review E* and the *Physical Review Letters*. The main reason for selecting the last journal in which respondents published was because that journal covered their area of interest. Other important factors were prompt publication, worldwide readership and a high Impact Factor.

As mentioned before physicists have been pioneers in the adoption of new technologies for scholarly communication and physics publishers have embraced electronic publications. In a discussion on the VPIEJ e-mail list, a Senior Associate from the American Astronomical Society claimed: "...paper journals (particularly ones with electronic counterparts) will fade away. In astronomy, this will happen within four years. By January 1998, 95% of the world's peer reviewed astronomical literature will be available online." (Ware, 2005:193). This prediction might not had come true by 1998, but certainly now in 2007 the vast majority of literature in physics and astronomy are available online.

The popularity of e-print archives which are open access sources of information has attracted attention of researchers with regard to the impact of open access papers and physics has always been a good subject to investigate this issue. One of the studies that investigated the citation to open access material in physics is Brown's (2001a) study. Using various sections of arXiv and the SPIRES-HEP database, she examined the citation rates of e-prints by

e-prints and concluded that e-prints have come of age in the literature of physics. Her results indicated that e-prints are used to a greater extent by physicists than previously measured and that e-prints have become an integral and valid component of the literature of physics. Moreover, her findings showed that High Energy Physics Experiment (hep-ex) had the highest citation rate at 14.5%, while Mathematical Physics (math-ph) had the lowest at 0.95%. She also used SciSearch database to analyze the citation pattern of journal articles to e-prints. In addition Brown stated that High Energy Physics Theory (hep-th) had the highest citation rate while Physics had the lowest at 0.07%. Citation rate by e-prints to e-prints was 20 times greater than the citation rate by journal articles to e-prints.

Another study by Hajjem, Harnad and Gingras (2005) performed a 10-year citation tracking of different fields and showed that open-access articles have a greater impact on research compared to closed-access articles. The research showed that physics had the highest ratio of citations of open-access articles to citations of closed-access articles published in the same issue of a given journal.

Lawal (2002) surveyed a random sample of 473 scholars from different fields and found that eighteen percent of the respondents used e-prints and 82% did not. Of those who used e-prints, 54.2% were in Physics/Astronomy, 27.7% were in Mathematics and Computer Science. Chemists used the e-prints the least due to publishers' policies. One hundred percent of those who utilized e-print archives also searched e-print archives but only 90.7% cited them in their articles while 9.3% do not. There were a large number of respondents in all areas that felt that e-print archives were not relevant to them. A relatively small number named technology constraints as a barrier to use. Forty per cent of physicists and astronomers replied that they would use e-print archives if the barriers were removed, 40% said no and 20 did not reply. Among physicists and astronomers 81.25% posted their paper to the e-print archive before publication and 17.5% after publication. 83.4% of those who posted their paper before publication published their article later. Seventy-two percent of respondents who used e-print archives said they did so for rapid and wider dissemination of information and fourteen percent said they do so for visibility and exposure. In terms of physicists' and astronomers' awareness of open access journals, a survey of 4000 journal authors around the world by Nicholas et al (2005b) showed that 62% of physicists and astronomers knew a lot or quite a lot about open access journals.

#### 2.7.4. Scholarly Communication and IT

Scientists use different communication channels for communicating with their colleagues and peers. Traditionally the scientific communication happens through written correspondences, meetings and publications (books and journals and so on). The application of information and

communication technology in the area of scientific communication has created new channels for communication (such as internet-based communications means) and transformed the traditional channels (such as journal publishing). As mentioned earlier physicists and astronomers have been among scientists who embraced use of IT and new communication technologies for scientific purposes. For example the survey by IoPP (Singleton, 1997) showed that almost everyone (98%) used e-mail and in mid-1995, just over half had Web access, and virtually all expected to have it by mid-1996.

Disciplinary differences in the utilisation of scholarly communication means and methods have been established by some studies. For instance, Walsh and Bayma (1996:869) showed clear differences in the use of computer-mediated communication for formal and informal communication across mathematics, physics, chemistry and experimental biology. They mentioned four structural factors affecting usage patterns of computer-mediated communication including: size of research field, market penetration, locus of critical information and degree of interdependence between research units, and technical limitations.

They further explained this by the hypothesis that market-buffered fields such as mathematics and physics would make extensive use of networked Information Communication Technologies for informal communication. On the other hand, market-penetrated fields such as chemistry and experimental biology would mainly utilise digital networks for formal communication and would make limited use of them for informal communication. Walsh and Bayma (1996) also highlighted the effect of work organization and organisation of disciplines on the uptake and use of computer-mediated communication technologies.

Besides the factors such as penetration by market, other factors are thought to influence communication patterns in different fields. For example, Kling and McKim (2000) listed four characteristics that lead scholars to shape the significance of peer reviewing and formal publication in their communication systems. They are research project costs, mutual visibility of on-going work in the field, degree of industrial integration, and degree of concentration of communication channels. They, for instance, explained that cost of the project may have several effects on the scholarly communication:

First, it may tend to increase collaboration, as it becomes difficult for any one researcher to mobilize the resources necessary to perform the research. Second, it may increase visibility of the work. Third, many specialized extramurally funded research institutes establish stronger on-going controls for publishing research results, even as working papers. Institutes as diverse as CERN and the RAND Corporation are known for their internal reviews. High-cost (multimillion dollar) research projects usually involve large scientific teams who may also subject their research reports to strong internal reviews, before publishing. Thus, a research

report of an experimental high-energy physics collaboration may have been read and reviewed by dozens of internal reviewers before it is made public. (Kling and McKim, 2000:1313).

Olson and Olson (2000) who investigated the factors affecting the selection of communication media, chose space physicists as their case study. They showed that many organisational contexts do not allow for the selection of appropriate communication media and that there is a tacit acceptance of existing technologies infrastructures that lacks a robust assessment of the available communication media. Furthermore, they maintained that the existing culture of a knowledge domain or other type of organisation will influence the adoption and use of computer-mediated communication technologies.

The indication of their research was that some communities are in more of a state of what they refer to as 'collaboration-technology readiness' than others. They showed that the existing culture of the space physicists, which was collaborative and encouraged the sharing of findings through informal modes of communication, lent itself readily to collaborating remotely via computer-mediated communication. Furthermore, they believed that the later adoption of the Web by the space physicists to communicate results and collaborate can be attributed to the familiarities of community members with less sophisticated computer-mediated communication technologies:

Later incarnations of the collaboration technology for the space physicists evolved with their general technical sophistication. When the Web became popular, others started putting relevant instruments online. Those who had already participated in the project began to demand access to those sites as well, and the entire project became Web based. As experience grew, they became more and differently collaboration-technology ready. The interface they have now would not likely have been accepted at the outset. (Olson and Olson, 2000:165)

Among subfields of physics and astronomy, high energy physics has attracted more attention due to its role in the development of e-print archives and also because it has probably clearer boundaries compared to the other subfields of physics. Traweek (1988) stated that the significant communication within high energy physics takes place by word of mouth in the form of casual interaction, or at intellectual social gatherings such as lectures and seminars. Knorr-Cetina (1999) described the role of gossip in the social organisation of high-energy physics and the development of interpersonal recognition. She developed the notion of 'technical gossip' that she described as an evaluative and personal discourse that interweaves 'report, commentary and assessment regarding technical objects and regarding the relevant behaviour of persons' (1999:205). She argues that this kind of gossip is a channel for the development of interpersonal recognition, because it reproduces a 'personalized ontology' that transcends organisational boundaries such as experiments, experimental groups and institutions.

Word of mouth and interpersonal communication appeared to be important in physics. Allen (1991:28) stated that oral communication has historically been a heavily used but seldom documented method for communication in physics. The most basic form of oral communication is informal personal interaction. It is quite natural that scientists who spend much of their time conducting research will tend to speak with colleagues about that research. This type of communication takes place wherever the scientist is, be it in an office, hallway, restaurant, library, gym, etc. As most of the formal restraints common to many other communication modes are lacking, this is an opportunity for physicists to let their creativity flow, (Kasperson, 1978) and to trade bits of scientific gossip (Traweek, 1988)

Hensman (1977) believed that even the formal framework (e.g. conference and seminar) often opens channels to more informal relationships between scientists. He reported that when physicists in the UK were asked to choose from a number of reasons why they attend conferences, most listed keeping in touch with current developments, meeting people with similar interests, and hearing papers presented in their field as being most important, though the list given to choose from was rather limited.

Fry (2006) used domain analytic approach to investigate scholarly research and information practices in different scholarly communities. Her case study included high-energy physics, corpus-based linguistics, and social cultural geography. Her research showed that Whitley's (2000) theory of mutual dependence and task uncertainty could be used as an explanatory framework in understanding similarities and difference in information practices across intellectual fields. Whitley's organizational theory of scientific fields and research governance integrates both cognitive and social aspects of scientific work. Fry and Thelwall (2006) explains that Whitley's theory

contributes to an explanatory framework for systematically comparing the cultural identity of scientific communities, their research objects, problems, goals, techniques, intellectual priorities, significance criteria, and reputations, across diverse knowledge domains. Whitley argues that many of the major differences between fields can be explained in terms of two dimensions of research practice: (1) the degree of mutual dependence between researchers in a field when making competent and significant contributions to the body of knowledge; and (2) the degree of task uncertainty in producing and evaluating knowledge claims (Fry and Thelwall, 2006).

The research which utilised interview as data collection tool showed that the way in which scholarly communities coordinate and control research problems, techniques, strategies, task outcomes and reputations will significantly influence the production and use of digital infrastructures and resources across fields. Furthermore the findings revealed that:

those fields that are non-hierarchical, loosely organised, intellectually pluralistic, with local variation in work organisation, are particularly likely to rely heavily

upon face-to-face informal communication for coordinating collaborative work, and will rely more heavily upon formal communication for community-wide dissemination of research and reputation building. The lack of centralised coordination and control in these fields will make it difficult for the scholarly community to systematically appropriate and develop digital infrastructures and resources in response to specific cultural needs. Often such fields have to work within externally imposed and developed digital infrastructures and resources. This then further compounds the problems associated with decentralised work practices e.g. appropriating modes of dissemination, lobbying for increased funds, and sustaining collaborative projects. (Fry, 2006:312).

Perhaps one of the characteristics of scholarly communication in physics and astronomy is that the move towards use of information technology and digital information services and systems is massively motivated and pushed by physicists and astronomers themselves. As Fluckiger (1989) argued years ago one reason why physicists have been at forefront of the database and networking technologies is the fact that much of the data collection done in physics research is carried out at research centres that are geographically distant from the researcher. This has created the development of state-of-the-art electronic networking.

### 2.7.5. Students' information-seeking behaviour

Since PhD students are part of the population of this study, it is logical to allocate a section to review of the literature of research on the students' information-seeking behaviour. Firstly, it must be said that PhD students' information-seeking tends to be different from undergraduates'. They are expected to have better information skills compared to undergraduates but relatively inadequate and less developed skills compared to faculty members'. Past research generally has confirmed the lower level of information seeking skills among students compared to faculty members (see for instance: Chu, 1995; Tillotson, 1995; Drabnstott and Weller, 1996; Bates, 1977; Hildreth, 1997; Tsai and Tsai, 2003; Scott and O'Sullivan, 2005) even at the level of research students (Zaporozhetz, 1987; Compton, 1989; Morner, 1993; Simon, 1995; Barry, 1997; Hess, 1999).

No research has been done on the information-seeking behaviour of research students in physics and astronomy and the only significant research involving physics postgraduate student is Brown's (1999b) study on their information literacy skills. She surveyed a group of graduate students in physical sciences at the University of Oklahoma. The study showed that the graduate students in physical sciences were information literate and experienced little library anxiety. However, they did not embrace the Internet as the key to the world of information.

Although as it was mentioned little related research has been done on research students in physics and astronomy, there have been some studies on research students in other subjects, some of which we mention here. With regard to the students' knowledge of source types (e.g. theses, students' supervisors, conference papers, and outside experts), Morner (1993) tested the library research skills (including source types) of 149 education doctoral students using a survey of 41 questions and found that, on average, students only answered approximately 50 percent of the questions correctly. For example, 60 percent of the students correctly answered the question 'which sources usually have the more scholarly, respected research?' The correct answer was 'refereed journals', but many, instead, chose dissertations, ERIC documents or text books. In terms of the relative importance students place on different types of sources, past research has identified the following source types important to students: academic journals, books (other than textbooks), dissertations, the students' supervisors, conference proceedings, experts outside the students' own institution, and bibliographies and references in journals and books. The findings of the longitudinal study by Chu and Law (2007) not only confirmed previous findings (such as Compton, 1989; Holland et al., 1991; Brown, 1999b; Cole, 2000; Chang and Perng, 2001) with regard to various source types (e.g. theses, students' supervisors, conference papers, and outside experts) important to research students but it also identifies some source types rarely mentioned in other articles as important (sources such as technical reports and patents).

Research by George et al (2006) on graduate students' (both doctoral and taught courses) information behaviour related to their process of inquiry and scholarly activities showed that they often begin with a meeting with professors who provide direction, recommend and provide resources. Other students help to shape graduate students' research activities, and university library personnel provide guidance in finding resources. The research found out that the Internet plays a major role, although students continue to use print resources. Convenience, lack of sophistication in finding and using resources and course requirements affect their information behaviour. Findings of the study varied across disciplines and between programmes.

Although it has been established that research students place different levels of importance on different sources of information, little research has been done on investigating students' reasons for using or not using certain source types. One of the few studies that has investigated this area is Farid's (1984). Farid (1984) showed that the two main reasons for PhD students using personal channels were: (1) the opportunity to discuss the contents of the material with the advisors and (2) material borrowed or references recommended which had been evaluated as to their importance for the students' research.



According to a study by King and Montgomery (2002) at Drexel University, doctoral students appeared to be more dependent on the library copies of journals than faculty members who had a greater number of personal subscriptions. Further, with regard to these preferences, it was found that faculty members often preferred to browse, whereas doctoral students in particular preferred to search. Observations made in other studies (such as Tenopir et al., 2003) also supported similar findings. However, the increasing access to electronic format of journals provided by libraries might have changed this situation.

Using both qualitative and quantitative data collection techniques, Chu and Law (2005) conducted a longitudinal study on twelve research students in engineering and education in order to find out about their development of information search expertise. Their study showed that many students were initially unfamiliar with many of the databases important to them and that the familiarity developed during the year contributed importantly to their development of information search expertise. The study also revealed that Students' perceived importance of databases changed over time.

In a later study Chu and Law (2007) identified three stages of students' information needs: general, specific and the most current. They showed that these various stages of information needs are closely linked to the students' research progress. Their study further showed that many students were initially unfamiliar with many source types important to their research, and the growth of knowledge of many of these sources has contributed to the students' development of information search expertise.

A review of the literature related to research students' information behaviour by Chu and Law (2007) led to identification of five aspects regarding source types used by research students that have been examined by researchers: (1) the kinds of sources they use (Brown, 1999b; Chang and Perng, 2001; Compton, 1989; Fabiano, 1996; Holland et al., 1991; Simon, 1995); (2) the extent to which they use these various source types (Chang and Perng, 2001; Compton, 1989; Fabiano, 1996; Farid, 1984); (3) students' perceived importance of or the extent of the usefulness of the sources (Brown, 1999b; Chang and Perng, 2001; Cole, 2000; Compton, 1989; Farid, 1984); (4) students' reasons for using or not using certain source types (Compton, 1989; Farid, 1984), and (5) students' knowledge of certain sources (Morner, 1993).

## **2.8. Summary**

This chapter presented a review of the literature in the field of information-seeking behaviour. The chapter started with introducing the concepts of information behaviour and information-seeking behaviour and their definitions. Then three of the main approaches to research in information behaviour research including cognitive approach, social approach and domain-

analytic approach were explained. The chapter then discussed the models of information-seeking behaviour, particularly Ellis's model. Information-seeking behaviour in the academic environment was discussed next as the population of this study are academics. The last main section of the chapter discussed different aspects of information-seeking behaviour and scholarly communication of physicist and astronomers. Studies related to the information-seeking behaviour of research students were also discussed. The review generally showed that although some issues such as preprint culture and scholarly communication have been well presented in the literature, there is a lack of up-to-date and comprehensive study on the information-seeking behaviour of physicists and astronomers.

## CHAPTER 3 - METHODOLOGY

### 3.1. Introduction

This chapter delineates the methodology that has been applied in this study. Issues concerning research approach, data collection methods, sampling and datasets, data analysis and limitations of the research are discussed. The chapter starts with a discussion on the qualitative and quantitative approaches to the research and continues with the description of different methods applied in the research and their specifications.

There are different methods concerning the study of information-seeking behaviour of users including questionnaire survey, interview, focus group, observation, transaction log analysis, citation studies, and diary recording. Each of these methods has their own advantages and disadvantages and no research method is entirely free from problems. However, the issue of choosing the most suitable method is an important and central issue in every single research. None of these methods is superior per se and the choice of methodology depends on several factors that a researcher should take into account. Aims and objectives of the research, research questions and limitations of the study are a few to mention.

Once the research aims and objectives have been specified and the research questions are matched to the research aims and objectives, the first major task is to decide about the appropriate research design and the methods of data collection. One of the significant factors that affect the decision of data collection method is the research questions as it is believed that the question of methodology is secondary to the research questions and content. Different questions require different methods to answer them. In some cases it is easy to identify the appropriate method for answering a question. However, there are situations where answering a question and achieving an objective could be done using different methods and approaches. These are the situations where the researcher may find it more difficult to decide which method or approach to use.

To decide which data collection method(s) is more suitable and efficient for answering the research questions, the researcher needs to think of the broader framework in which the research methods fit and choose the right approach to the research. This framework that could

be referred to as research approach, research strategy or research paradigm<sup>5</sup> can traditionally be of two main types: qualitative or quantitative.

### 3.2. Qualitative approach vs quantitative approach

Studies with a quantitative approach deal with numbers and figures (i.e. quantities) while studies with a qualitative approach deal with descriptions of concepts and perceptions. The results obtained from the two approaches can be very different. Quantitative studies seek findings that can be used to make generalisations across the field of research. Although qualitative research can lead to generalisations, its prime function is to help interpret phenomena within 'real-life' contexts (Mansourian, 2006:48). The two approaches, i.e. the qualitative and the quantitative, have important differences. Table 3.1 summarises some of the differences between these two approaches.

**Table 3.1: Differences between quantitative and qualitative approaches based on Hammersley (1992) and Punch (2005).**

Quantitative	Qualitative
Its methods in general are more one-dimensional and less variable, hence more easily replicable	Its methods are less formalised and more multi-dimensional and diverse and less replicable, hence greater flexibility
Conceptualizes reality in terms of variables and relationships between them	It deals more with cases, and is sensitive to context and process
Typically has larger samples, guided by probabilistic considerations	Typically has smaller samples, guided by theoretical considerations
Uses numbers	Uses words
Concerned with behaviour	Focuses on meaning
Is more concerned with the deductive testing of hypotheses and theories	Is more concerned with exploring a topic, and with inductively generating hypotheses and theories
Is more powerful for generalisation	Lacks the power of generalisation

Generally in social and behavioural sciences there has been a debate known as quantitative versus qualitative debate during the last three decades prior to the twenty-first century. This debate became increasingly unproductive during 1980s and early 1990s before it

<sup>5</sup> The terms research strategies, research paradigms and research approaches have been used interchangeably in the methodological literature for discussing qualitative and quantitative approaches. The researcher uses 'approach' in this work as it seems to be more common.

eventually ended. The debate, known as paradigm war, has been positive for research development in many fields as most researchers now use whatever method is appropriate for their studies, instead of relying on one method exclusively (Tashakkori and Teddlie, 1998).

One of the outcomes of the debate was the belief, mainly by pragmatists, that qualitative and quantitative methods are indeed compatible (Tashakkori and Teddlie, 1998:3). Punch (2005:235) also pointed out that some of the aforementioned distinctions between the two approaches are based on stereotyping. For example, quantitative approach can also be used for exploring an area and for generating hypotheses. What can be concluded from the characteristics of the two approaches and the outcome of the Qualitative vs Quantitative debate is that we cannot find out everything we might want to know using only one approach and we can often increase the scope, depth and power of research by combining the two approaches. An intensive review of literature by Tashakkori and Teddlie (1998:x) ended with the approval that mixed model studies are the growing trend in the social and behavioural sciences. Brannen (2004) also highlighted the pressures towards convergence of the two paradigms that suggest a move away from the separate paradigms model. These pressures are caused by causes, among them is what Bernstein (2000) considered as the erosion of boundaries around social science disciplines.

### **3.3. Qualitative and quantitative research in LIS**

The same trend, which was described above, holds true for research in LIS and more particularly research on information behaviour that is the concern of this study. Quantitative (or positivist) approach, which is considered as traditional approach, has been the popular approach in information behaviour studies. Generally, in the field of information science, approaches to data gathering and analysis have predominantly been quantitative (Fry, 2003:65). Questionnaire survey with close-ended questions (and other quantitative methods such as log analysis) is probably the most used method in information behaviour studies. Studies which are based on quantitative approach have made a valuable contribution in our knowledge of information behaviour of users.

However, there have been some criticisms against mere quantitative approach to information behaviour research. For example, Wilson stated that:

...in the positivist tradition, quantitative research methods were adopted that were inappropriate to the study of human behaviour: many things were counted, from the number of visits to libraries, to the number of personal subscriptions to journals and the number of items cited in papers. Very little of this counting revealed insights of value for the development of theory or, indeed, of practice (Wilson, 1999:250).

It must be said that this sort of criticism is by no means meant to undermine or invalidate the value of quantitative research. As a matter of fact these criticisms rather indicate the lack of qualitative research, which is an appropriate approach for conceptual enrichment of the field. During the last decade the situation has changed and as Bawden stated

it is no exaggeration to say that qualitative research is now the most common approach for information research involving information seeking and use, alone or in conjunction with quantitative methods (2006:676).

The development of many of the existing theories and models in the field of information behaviour is owed to qualitative research and there have been some rich qualitative studies such as Ellis (1989); Ellis, Cox and Hall (1993); Kling and McKim (2000); Brockman et al. (2001) and Ford & Mansourian (2006).

Onwuegbuzie, Jiao and Bostick (2004:83) believe that in library and information science as in the case for social and behavioural sciences, all research questions can be addressed using either qualitative, quantitative, or mixed-methodological research techniques.

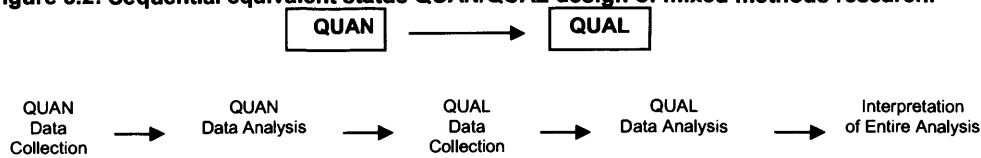
### 3.4. Mixed methods

As argued above, there is a third to the duality of qualitative and quantitative. Mixed-methods that is also referred to as multi-method, mixed methodology, methodological mixes, and triangulation is an approach to research that contains elements of both qualitative and quantitative approaches. Tashakkori and Teddlie (1998:5) listed three main mixed method designs:

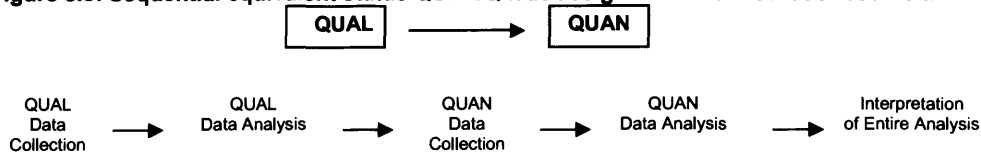
- Equivalent status design: sequential (QUAN/QUAL and QUAL/QUAN) and parallel/simultaneous (QUAN + QUAL and QUAL + QUAN); this is the case where a researcher uses both approaches about equally to understand the phenomenon under study. The use of the two approaches can be sequentially in terms of time or simultaneously. Figure 3.1, Figure 3.2 and Figure 3.3 illustrate the different types of equivalent status design.

**Figure 3.1: Simultaneous equivalent status QUAL/QUAN design of mixed methods research (Creswell, 2003, p. 214).**

**Figure 3.2: Sequential equivalent status QUAN/QUAL design of mixed methods research.**



**Figure 3.3: Sequential equivalent status QUAL/QUAN design of mixed methods research.**



- Dominant-less dominant designs: sequential (QUAN/qual and QUAL/quan) and parallel/simultaneous (QUAN + qual and QUAL + quan); this is the case where both approaches are used by the researcher but one of them is dominant. The use of the two approaches can be sequentially in terms of time or simultaneously. Figure 3.4 relates.

**Figure 3.4: Dominant- less dominant mixed methods research (Creswell, 2003, p. 214).**

- Designs with multilevel use of approaches. These are studies in which data from more than one level of organisations or groups are used to reach more comprehensive inferences regarding behaviours or events.

With a different point of view, Hammersley (1996) (as cited in Brannen, 2004:314) suggested a tripartite classification of the ways in which researchers employ different types of data in the processes of interpreting their data:

- Triangulation, where one type of data (usually quantitative) is used to corroborate another type of data (typically qualitative), as when theoretical insights are derived from one type of data which are also put to the test on another dataset.
- Facilitation, where collecting one type of data facilitates the collection of another type of data; for example, when qualitative interviewing methods are first employed in preliminary pilot work in order to help design a large-scale pre-coded survey.

- Complementarity, when two different sets of data are employed to address different but complementary aspects of an investigation; for example, qualitative data are used to understand social processes while quantitative data are employed to examine associations and their statistical generalisability to parent populations.

Using mixed methods can be very time consuming as it employs the use of different levels of analysis imposed by the different methods. However, there are several advantages to mixed methods. The validation of research data and methods through triangulation, encouragement of creativity that could stimulate further work and expansion of the scope of the study are among benefits of using a combination of methods (Tashakkori and Teddlie, 1998). Greene and colleagues (1989) added complementarity, development and initiation to the list of triangulations' benefits; and Rocco and colleagues (2003) highlighted the increased opportunity for exploratory inductive process that begins with empirical evidence of the particular and proceeds with to a level of abstracting/theorizing/generalizing and the confirmatory deductive process of hypothesis testing of theories.

### 3.5. Research design

There are different uses of the term 'research design' in the literature. 'All the issues involved in planning and executing a research project' is the definition used by Miller (1991) at the most general level. Bryman (2001:29) maintains that research design provides a framework for the collection and analysis of data. The choice of research design reflects decisions about the priority being given to a range of dimensions of the research process. These may include: expressing connection between variables, and generalisation to a larger body of data. Oppenheim (1992) points out that the term 'research design' refers to the basic plan or strategy of the research, and the logic behind it, which makes it possible to draw valid general conclusions from it. The choice of research design is concerned with making the study problems researchable by setting up the study in a way that produces specific answers to specific questions. Punch (2005) believes that research design is the basic plan for a piece of research, and includes four main ideas: Strategy, conceptual framework, who or what will be studied, and tools and procedures for data collection. Therefore the data will be collected and analysed by answering these four questions: following what strategy? Within what framework? From whom? and how? Design sits between research questions and the data, showing how the research questions will be connected to the data, and what tools and procedures to use in answering them. On the other hand, research methods are techniques for collecting data. It can be a specific instrument such as a self completion questionnaire or a structured interview, or a participant observation used for data collection (Bryman, 2001).



As described, different possible scenarios can be applied in a mixed methods research. This study applies the 'sequential dominant - less dominant design (QUAL/quant)' of mixed methods. The study uses sequential mixed methods in order to gain a more robust set of data and achieve a more comprehensive understanding of them.

Creswell (1995, Cited in Tashakkori & Teddlie, 1998:46) called this type of research design a two-phase design and pointed out that this research design is popular with graduate students and novice researchers wishing to use both approaches in their research while trying to stay away from the complexity of using both approaches simultaneously.

The current research is based on the QUAL/quant sequence meaning that the investigator starts with qualitative data collection on a relatively unexplored area or with a neutral approach, and then uses the results to design a subsequent quantitative phase of the study. The dominant approach is qualitative and the less dominant one is quantitative. The quantitative method is used in this study to clarify and replicate some of the results of the qualitative study as well as investigating some other aspects of the information-seeking behaviour of the participants that are not covered by qualitative part of the study.

The reason for more focus on the qualitative approach in the current research is that the main aim of the research is to derive behavioural patterns of information seeking patterns of academic scientists. Gorman & Clayton in their book on qualitative research in the field of library and information studies defined qualitative research as:

[A] process of enquiry that draws data from the context in which events occur, in an attempt to describe these occurrences, as a means of determining the process in which events are embedded and the perspectives of those participating in the events, using induction to derive possible explanations based on observed phenomena. (2005:3)

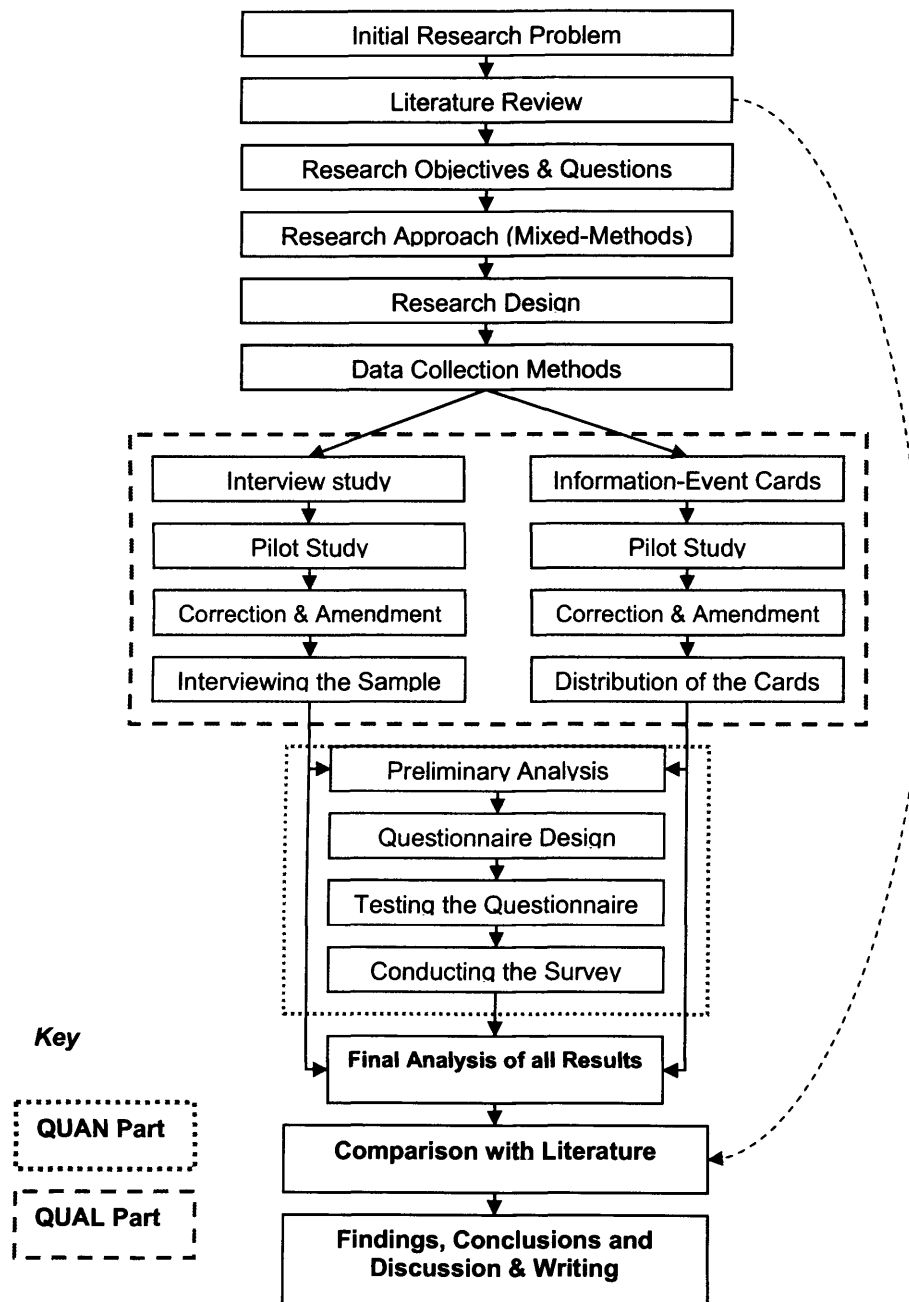
According to them, the purpose of qualitative research can be interpretive, understanding participant perspective, or contextualisation and the approach can be pattern seeking or theory generating among others (Gorman & Clayton, 2005:4).

The research process started with neutral enquiry into information-seeking behaviour of physicists and astronomers. The first phase of the research was qualitative data collection by means of interviewing. Once the preliminary results of the qualitative data were obtained, quantitative data collection was conducted to clarify, replicate and triangulate some of the outcomes of the qualitative data. Eventually a comprehensive analysis was done on all the qualitative and quantitative data and conclusions were drawn.

It should be mentioned that the methodology applied in this study is one of the strong points of the study for it takes advantage of several data collection techniques to reduce the

effect of each techniques pitfalls and mixed both qualitative and quantitative data to gain a picture as holistic as possible. Figure 3.5 is a flowchart diagram that depicts the stages and the process of this study.

Figure 3.5: The stage of the research process.



### 3.6. The research population

The research population of this study is the Department of Physics and Astronomy at University College London. The reason for studying physicists and astronomers was explained earlier in the first chapter. Here the specifications of the general research population are discussed.

#### 3.6.1. UCL Department of Physics and Astronomy

The department of physics and astronomy at UCL is one of the largest and oldest departments in its field in the United Kingdom. Its history can be traced back to early 19th century (Fox, 2007). It was rated at 5 in the last Research Assessment Exercise (RAE 2001). Hagstrom (1970:108-9) pointed out that large and high prestige departments might be expected to be centres of communications and all members of these departments, including those with little personal prestige, might benefit from this position. This is because, according to him, most university scientists communicate more with their departmental colleagues than with the others, and they are often introduced to the work of scientists in other institutions by their departmental colleagues.

UCL department is a research oriented department that at the time of this research (2005-06) had about 150 academic and research staff and more than 100 postgraduate students. The department's research performance has been improving constantly during the last 30 years. Figure 3.6 and Figure 3.7<sup>6</sup> generated using ISI Web of Science<sup>SM</sup> Analyze Tools show that the number of publications published by physicists at UCL and the number of citations their publications have received have been in steady increase since early 1990s.

---

<sup>6</sup> The diagrams were generated ISI. A search was run for all materials published by UCL and the results were refined by document type (was restricted to articles) and subject area (which was restricted to Astronomy and Astrophysics, Nuclear physics, Atomic, Molecular and chemical physics, and particle physics. The result was 3,055 articles published between 1972 and 2007 that have received 46,632 citations between 1972 and 2007 (up to 28 September). It is appreciated that these are not all of the subjects covered by research activities in the UCL department of Physics and Astronomy and also some of these articles might have been published by UCL staff in other departments other than Dept. of Physics and Astronomy. However, it still gives a good image of the UCL output and performance in the field of physics and astronomy of the purpose of this study.

Figure 3.6: Number of articles published by UCL in physics and astronomy (ISI data).

### Published Items in Each Year

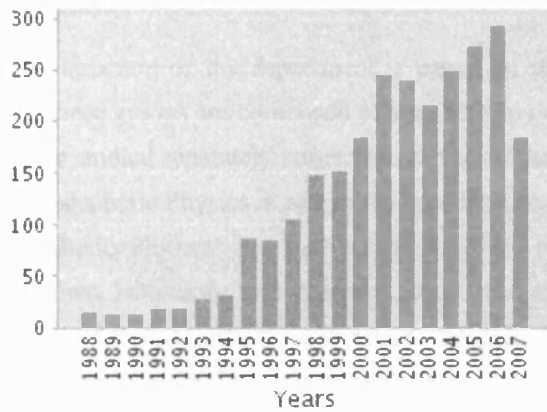
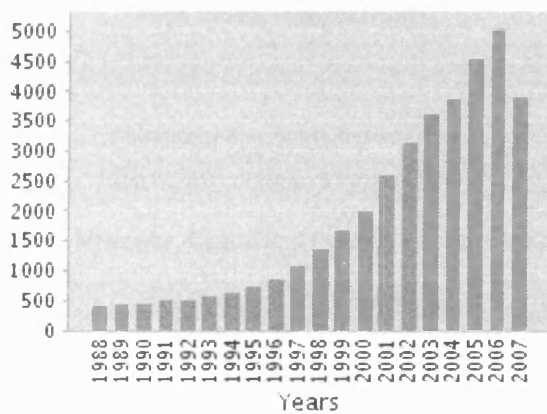


Figure 3.7: Number of citations received by articles published by UCL in physics and astronomy (ISI data).

### Citations in Each Year



The department consisted of four research areas and contributed to six research centres that each had their own researchers. The four research areas were:

- Astronomy, Astrophysics and Atmospheric Physics
- Atomic, Molecular, Optical and Positron Physics
- Condensed Matter and Materials Physics

- High Energy Physics

There was another research group, nominally ‘Image Processing’ which practically was not active and most had been merged with the Department of Computer Science and therefore, was not included in the research.

Although the main structure of the department is based on the four aforementioned research groups, some of these groups are composed of smaller research groups that are quite characteristics and could be studied separately rather than as part of the bigger research group. For instance, although Atmospheric Physics is part of the broader research group ‘Astronomy, Astrophysics and Atmospheric Physics’ the first research group, it is a quite distinctive research group with its own laboratory and research areas that are not very related to astronomy. Or in the case of the second research group ‘Atomic, Molecular, Optical and Positron Physics’, two subgroups of ‘Optical Science Laboratory’ and ‘Theoretical Molecular Physics’ could be separated as two distinctive groups. Therefore the researcher has decided to consider the following seven research groups (Table 3.2) as the research areas in the department and units of analysis in this study wherever appropriate.

**Table 3.2: Research groups in the Department of Physics and Astronomy.**

Atmospheric Physics (AP)
High Energy Physics (HEP)
Condensed Matter and Materials Physics (CMMP)
Astronomy and Astrophysics (AA)
Theoretical Molecular Physics (TMP)
Atomic, Molecular, Optical and Positron Physics (AMOP)
Optical Science Laboratory (OSL)

### 3.7. Research methods used

Four different methods were used in this study including desk research, semi-structured interview, information-event cards and questionnaire survey.

The study entailed some desk research that was carried out mainly in the beginning of the study and also during the course of the study. The main part of this study was based on the interviews and the information-event cards. After the preliminary analysis of the findings of this phase, a questionnaire survey was set in order to answer some of the questions that were raised by the results of the interviews and also in order to collect some complementary data. In

the following section, each of the used research methods will be described. Table 3.3 shows the research methods and the time-line of the study.

**Table 3.3: Research Methods applied in the study.**

<b>Method</b>	<b>Time</b>
Desk research	Prior to and throughout the study
Interview, Pilot study	August – September 2005
Interview, main study	October 05 – April 06
Information- event cards	October 05 – April 06
Questionnaire survey	April – June 06

### 3.7.1. Desk Research

The research entailed some desk research. Desk research included literature review as well as obtaining existing data and information about the subjects of the research, the research population (the Department) and so on.

In order to identify the related literature for the literature review, different strategies were applied. Extensive searches were carried out in online version of the database 'Library and Information Science Abstracts (LISA)' to find any work pertinent to information-seeking behaviour of scholars, particularly physicist and astronomers, and other issues that played a part in this research such as information-seeking behaviour of research students. Although LISA database is considered a very exhaustive database for the LIS area, searches were also conducted in Google Scholar and Web of Science to make sure that the researcher does not miss out any relevant works in case they were not indexed in LISA. The advantages of databases such as Google Scholar and Web of Science are that they do not limit the search to the literature of a specific field. They also provide citation linking and tracking facilities. However, the search strategy was just a point to start for finding the relevant literature. The researcher identified a considerable part of the literature he read through looking up the references at the end of key relevant papers and also searching for the works that might have cited the key works using Google Scholar and the Web of Science. The researcher has been also subscribed to the table of content email alerts of most of LIS journals since end of 2003 and he has kept an eye on the literature spotting any paper that might have had a link to his research interests. It is worth mentioning that what is presented here is a small part of the bulk of the literature that the researcher read to conduct the research.

The desk research also included obtaining research or raw data about for example characteristics of different subfields of physics and astronomy, the Department of Physics and Astronomy at UCL that was the research population, and participants in the interviews.

### 3.7.2. Interview

The main part of this study is based on in-depth, semi-structured open-ended interviews that were conducted in a face-to-face manner. Interview - like any other technique- has its own strengths and weaknesses. Gorman and Clayton pointed at some of the advantages and disadvantages of interview. They mentioned mutual exploration, investigation of causation and personal contact as advantages of interview. Moreover they stated that:

interviewing allows both parties to explore meaning of questions posed and answers proffered, and to resolve any ambiguities. Open-ended questions, in particular, may lead to unexpected insights. The other advantage is that interview can enable a researcher to explore causation. That is, to enquire into why individuals or organisation behave in the way that they do, something that most quantitative research cannot really answer (Gorman & Clayton, 2005:125).

The latter particularly is one of the objectives of this research and hence is the adoption of interview as the main data collection method.

Interviews also facilitate the collection of a large quantity of rich data in a relatively short space of time. Nicholas (2000:111) also praised open-ended interview as the 'real star' among other methods of investigating information needs (Nicholas, 2000). However, interview is not free of difficulties and problems. It can be costly, very time-intensive, uncritical, personal, and open to bias. Interview can have some other negative aspects including lack of selectivity, which means that sorting out the important points from a large quantity of data can be difficult, and may raise questions about selective reporting. Verbal data are also particularly susceptible to errors in interpretation (Gorman and Clayton, 2005:126).

The inclusion of interview technique and qualitative approach in the research is because of the research's objectives and questions. A research by Brannen and Moss (1991:19) proved that qualitative data were particularly useful to establish patterns of behaviour. It is known that objectives in behavioural (and social sciences) can be categorised under five main categories, which are exploration, description, explanation, prediction, and influence (Johnson and Christensen, 2000).

The current research is exploratory. This exploratory nature of the research requires rather a qualitative approach to research design. According to Onwuegbuzie, Jiao and Bostick:

...in qualitative research, the research problem needs to be explored because little information exists on the topic. The variables are largely unknown, and the researcher wants to focus on the context that may shape the understanding of the

phenomenon being studied. In many qualitative studies, a theory base does not guide the study because those available are inadequate, incomplete, or simply missing. Qualitative researchers tend to use a more inductive form of logic. This type of reasoning provides context-bound information leading to patterns or theories that help to explain a phenomenon. Broadly speaking, qualitative research is the collection, analysis and interpretation of words and observations. Qualitative studies are conducted in order to increase insights and generate meaning for whole situations and abstract concepts. The outcome of qualitative research is the development or expansion of theory. The quantitative approach is regarded commonly as the traditional or the positivist approach (2004:86).

### 3.7.2.1. Sampling

In general, while in quantitative studies we require a sample which should be statistically representative of the whole research population, in qualitative research the sample does not need to satisfy this requirement. Gobo (2004:435) in a long article discussed this issue. He stated that studies conducted based on representative and non-representative samples are both legitimate because there are two kinds of generalisations: a generalisation about a specific group or population (which aims at estimating the distribution in a population) and a generalisation about the nature of a process. The requirements are completely different in the two cases. While the first generalisation is based on statistical logic, the latter is based on theoretical considerations.

However, as this research does intradisciplinary comparison, it is necessary to include representatives from all of the main research areas of the department in the sample. It was also necessary for the sample to include academics with different status and levels of skills including professors, researchers, lecturers, and research students. The department consists of about 250 research and academic staff and research students. The staff includes professors, readers, lecturers, researchers, pot-doctoral research fellows, and research students. The sample included interviewees with all of these different statuses. See chapter four for demographics of the sample.

The sampling method for the interviewing stage of the current research was purposive stratified non-random sampling. In describing this method, Tashakkori and Teddlie (1998:76) stated that it is similar to stratified sampling but in a non-random purposive, convenient manner. Case or cases are selected non-randomly (volunteer, available, and so on) from each subgroup of the population under study. In sociological research, this is also known as quota sampling.

The sample was also self-selected in a way because the participation was voluntarily. This was partly because the researcher had no authority to select people and make them collaborate and had to rely on volunteer participation.



The research did not include any undergraduate or masters students. The information-seeking behaviour of undergraduate students is expected to differ from faculty's due to a number of factors. Undergraduate students' information-seeking skills are not as well developed as faculty's. They also have different information needs, and seek information in a different context; they normally seek information to address the imposed questions by lecturers rather than self-selected questions. Masters' students could be considered in a transition period where they gradually get acquainted with the research process. However, the research projects they conduct for their dissertations are not as dense as the ones done by PhD students and faculty members. PhD students, who are mainly involved in research activities, are more integrated in their departmental information environment and their communication and information-seeking behaviour is expected to be more similar to the faculty's. There have been some studies on both taught graduate (masters) and undergraduate students mainly as users of library services (see Abdoulaye, 2002; Barrett, 2005; Callinan, 2005; Fidzani, 1998; Jankowska, Hertel & Young, 2006; Majid & Tan, 2002; Washington-Hoagland & Clougherty, 2002; Whitmire, 2002). The studies revealed differences (and sometimes similarities) between students' and faculties' information behaviour and also between undergraduates' and graduates' information behaviour. For example a study on graduate students in humanities showed that although there were substantial areas of overlap, the model of graduate student information-seeking behaviour that emerges from the study was not a clear reflection of either faculty or undergraduate models (Barrett, 2005).

Study of research (PhD) students' information seeking-behaviour could shed some light on the way future scientists develop their information seeking skills. Apart from Brown's study (1999b) on the information literacy of graduate students in physical sciences (chemistry, physics and mathematics), no study has particularly investigated communication and information seeking activities of PhD students in physics and astronomy.

### **3.7.2.2. Email interviewing**

Although the researcher did his best to conduct face-to-face interviews, but because three of the research students were not accessible due to the physical distance, the researcher used email interviewing technique to interview them. One of the students was writing up and far from London and the other two students were based in laboratories again far from London and it was not possible for the researcher to meet them in person. Email interviewing, though far a younger method compared to phone or face-to-face interviewing, is a well-established interviewing technique. For example Lokman Meho conducted his PhD project in information behaviour research using only this method (Meho and Tibbo, 2003). Meho (2006) in an article in *Journal of the American Society for Information Science and Technology* has discussed the

advantages and challenges of email interviewing. He concluded that while a mixed mode interviewing strategy should be considered when possible, e-mail interviewing can be in many cases a viable alternative to face-to-face and telephone interviewing. This is exactly what has been done in this study.

### **3.7.2.3. Pilot interview study**

It was vital for the researcher to do a pilot study for a few reasons. First of all, it was a way of ensuring the feasibility of the study. It was an opportunity for the researcher to develop his interviewing skills and become prepared for the main study. And it also helped him develop and clarify the interview protocol and questions.

Seven volunteer PhD students in the Department of Physics and Astronomy were interviewed for the pilot study between August to September 2005. The aim of the pilot study was to find out about information-seeking behaviour of PhD students in physics and astronomy and also to know about similarities and differences between the behaviour of students in different research groups.

### **3.7.2.4. Interview data collection procedure**

*Prior to the interviews:* Permission was obtained from the head of the Department of Physics and Astronomy at UCL at the time, so that the researcher was allowed to approach the staff and students in the department and seek their co-operation for the study (see Appendices 1 & 2).

The Department has a weekly newsletter that is distributed via email. A short note (Appendix 3) about the research was published in the newsletter a month before the start of interviewing stage in order to inform people in the department.

Personalized friendly emails (Appendix 4) were sent to the participants addressing them in a friendly manner (forename in the case of students and Dear Dr. or Professor for the rest) and asking them to participate voluntarily in the study. The subject and aim of the study, confidentiality of the data, and the permission from the head of the Physics Department for undertaking the study were mentioned in the emails.

The researcher preferred to start the interviewing stage with PhD students. The reason was twofold: firstly, because the researcher himself was a student it was easier to make a rapport with the interviewees and conduct the interviews while gaining experience for conducting more complex interviews with faculty members. Secondly, because PhD students' activities are more focused (their dissertation research), the interview would have been more straightforward compared to faculty members who are involved in a wider range of teaching

and researching activities. Therefore, the researcher opted for starting with more straightforward interviews first.

*Time and place:* Current research's participants were interviewed during October 2005 - April 2006. The interviews were carried out in mutually agreed time and place at the convenience of the interviewees, in most of the cases, in the common rooms in the Department of Physics and Astronomy, or in the participants' offices. The interviews were conducted face-to-face by the researcher. The interviews had variable length from about 30 to 70 minutes with an average of about 44 minutes. The average interview with students was shorter than the average interview with members of staff.

*Conducting the interview:* At the start of each interview, the interviewees were briefed on the main purpose of the research and particularly on the objectives of interview. It is worth mentioning that they had already been informed briefly about the research via the invitation emails. However, this was done as a reminder and also to make the atmosphere friendlier and give the interviewees an opportunity to ask any questions about any issues concerning the research or the interviews. The researcher also reassured them that whatever they would say in the course of the interview will remain anonymous and their name would not appear in any publication whatsoever.

During interview sessions the researcher employed the main skills of interviews including clarification, active listening, paraphrasing, and body language and so on to encourage the participants to talk about the questions in more details.

The interviews were recorded after obtaining permission from the interviewees for doing so. A digital voice recorder was used to record all interview sessions from beginning to the end. Digital voice recorders produce computerised audio files which can be transferred to a personal computer. These files may be played several times without any damage to the voice quality. The advantage of employing this kind of voice recorder in comparison to the ordinary recorders goes back to the quality of recorded voice and the easiness of storage, management, and transcription. Using digital recorder also makes the transcription work easier.

*After interviews:* Krueger stated 'transcript-based analysis is the most rigorous and time intensive of the choices' (1994:143). To conduct a rigorous analysis, all interviews were transcribed by the researcher. Transcribing simply means to type the notes or interview into a word processor. This, though necessary for analysis, makes the information much more accessible and easier to analyse. Transcription is a very time consuming and painstaking work (Poland, 2003). However, transcription of interviews by the researcher himself has many

advantages over hiring a transcriber. It helps the researcher get acquainted with the dataset and get prepared for the analysis stage (Mansourian, 2006).

After completion of each transcription the transcribed text was sent to the participants to provide them with an opportunity to make any changes and confirm the accuracy of the transcription. It was very helpful because some of the participants made some comments on the text and provided the researcher with further explanations. This was a good way to have confirmation of the accuracy of transcriptions, and an opportunity to follow up any points that the interviewees wanted to explain more. Moreover, in terms of research ethics the participants had a second chance to change what they had said or remove any part of it or even stand down of the research.

*Presenting the interview data:* Having interviewed 56 people, it is self-evident that it would not be possible, nor would it be desirable, to present a high percentage of the data in the dissertation. Therefore, the researcher was very selective in presenting quotations from the interviews. There were a few points that were taken into account for presenting the interview data in the dissertation.

As spoken language is somehow different from writing language, a direct quote from a dialogue might seem slightly disjointed because of existence of some repetitive and unfinished sentences or unnecessary catch phrases such as 'you know', 'I mean', 'you see' etc. In order to present interview quotes in a brief, concise and fluent way, and for the sake of succinctness in the thesis, in some quotations the researcher had to remove some of unnecessary or repetitive words. The removed parts of the quotes are shown by three dots (...) and also sometimes the researchers added a few words to some quotes in brackets to show that these words are for more explanations.

The interviewees happened to refer to other people (friends, colleagues etc.) by name at times. In those cases the names or a few other words were removed from the quotation in order to keep the presented data absolutely anonymous.

### **3.7.2.5. Interview protocol**

A protocol was used to conduct the interviews (see Appendix 5). The interviews started with asking the respondents about their background, research group membership, the nature of the research activities they were involved in and the type of research they normally conducted.

The interviewees were asked about the way they would approach to a new research project. In the case of students the focus of this section was on the way they approached their PhD research project. One of the things that were meant to be found in this part of the

interview was any patterns or trends in the way research was carried out in different research areas in physics and astronomy.

Then the interview was continued covering the following main areas: different techniques and methods used by interviewees for keeping abreast with the developments of their research fields; use of different information resources and the purpose of using them; means of obtaining articles, reading habits and managing obtained articles; publishing patterns and interaction of users with e-print archives; the role and importance of conferences and seminars in information seeking and conducting research; culture and mechanisms of communication within the department, and role of colleagues in information seeking and communication activities; problems or difficulties that they might have faced in their research activities in terms of information seeking, access to material and or anything related to your information needs; and changes in information seeking activities and trends over time.

### **3.7.2.6. Analysing interviews**

Coding the textual data and extracting meaning from them is one of the main stages in a qualitative research. Coding is allocating specific meaning to the data and organising the data into categories and classes and making the dataset manageable, so the data can be interpreted systematically.

To put it simply, qualitative analysis is the way in which researchers go about making sense of the data they have collected, so that they can communicate their findings to others via reports, books and articles. There are numerous ways of going about making sense of data including textual analysis, content analysis, conversational analysis, biographical analysis and so forth. The choice of the analysis method is normally aligned with the research method, theoretical framework, disciplinary area and topic. However, there are no strict rules, which have to be followed in qualitative analysis (Williamson and Bow, 2002) and this flexibility is one of the strength of qualitative research.

This research used a very common technique of analysis, which is using categories that allows researchers to code and retrieve. According to Richards and Richards:

the code and retrieve process consists of labelling passages of the data according to what they are about or other content of interest in them (coding or indexing), then providing a way of collecting identically labelled passages (retrieving) (1998:214).

An initial broad set of categories and codes were used for the preliminary data analysis as suggested by Gorman and Clayton (2005). These categories were based on the themes of the interviews. This preliminary data analysis allowed the researcher to begin examining data carefully and thoughtfully, and start breaking them into smaller units for more detailed

analysis. The data were analysed in a nonlinear manner and went through several rounds of reading and every time the analysis and categories became more detailed.

The analysis process can be sometimes hastened, as well as enhanced, with the aid of computer technology. At the time of analysing the data for this study, there were some software packages for computer-assisted analysis of qualitative data. Probably the three most popular packages were QSR NVivo and QSR N6<sup>7</sup> (formerly know as NUD\*IST), and ATLAS.ti<sup>8</sup>. Use of these software packages has advantages such as easier retrieval of the data, pattern analysis, conceptual mapping, and data visualisation (Seale, 2003). The researcher had the experience in using one (N6) of these software packages due to collaboration in some other research projects (see for instance Nicholas, Jamali and Rowlands, 2006; Nicholas et al., 2005c). However, he decided not to use any software for the analysis and manually analyse the data. QSR N6 software has restriction for the type of data entry and it only accepts plain text files. The power of this software is in that it helps researcher quantify the data by giving percentage and numbers of text units (sentences, paragraphs and so on) that have been allocated specific codes. This software package was not considered helpful for this research as quantification had no application in this study. The power of ATLAS.ti is its data visualisation facilities. But this was not the specifically the objective the researcher either. The interface of this software is not very user-friendly. The best software that could have been used for the research was QSR NVivo. It supports different formats for data entry, it has a user-friendly interface. Moreover, it does not have the restriction that N6 has in terms of the text units. However, after considering all of the factors and elements including cost-efficiency consideration and the fact that the software package was not available to the researcher, the researcher decided not to use any software and do the analysis manually.

### **3.7.2.7. Memo writing**

A memo simply is a document that is used to write ideas or information about interviews or categories (Williamson and Bow, 2002). As Gorman and Clayton (2005:218) put it, writing memos is a way of stand back from data immersion. It is essential for an investigator to capture reflections and insights as they occur. Memo writing is a good way of data reduction and they can be basis of the final report. Memos help researcher to record reflections and insights as they occur.

---

<sup>7</sup> [www.qsrinternational.com](http://www.qsrinternational.com)

<sup>8</sup> [www.atlasti.com](http://www.atlasti.com)

### 3.7.3. Information-event card

Since this study has an inductive approach, it is important to get a rich and comprehensive picture of the information seeking activities of users (Lee & Baskerville, 2003; Miles & Huberman, 1994). Spradley (1980) recommends highly intrusive techniques as participant observation for this type of research. However, information seeking is an intangible process and not performed at one point in time, it is impractical, if not impossible, to use this technique successfully. Therefore, the researcher decided to use a kind of critical incident technique (CIT).

CIT has its roots as a scientific method in psychological research in the USA army in 1940s, however, its background is as old as early 19th century (Flanagan, 1954). CIT is recognised as a valid, reliable, and effective method for gathering rich qualitative data for a variety of purposes, including the analysis of information behaviour. Fisher and Oulton (1999) discussed its application in library and information management research. Urquhart et al (2003) also discussed CIT and listed numerous studies in information behaviour that applied this method (Fisher & Oulton, 1999; Urquhart et al., 2003). CIT is considered to give one of the most accurate and reliable retrospective reports of processes in practice.

CIT implies the kind of information collected and not the tool used for data collection. Therefore different tools can be used for collecting critical incident information. For example interviews can be conducted to ask participants about incidents or CIT-related questions can be asked through a questionnaire survey. The other possibility - that was applied in this study - is information-event card.

Information-event card is a self-report technique by which a user records the details of a single information incident. The researcher has borrowed the name 'information-event card' from Rubenstein et al (1970). They used a similar tool to collect self-kept record of information seeking activities of medical researches. An information event is an occasion when users need and look for some information. In this research, information-event cards were used as a complementary method of data collection and not as a substitute for interview or questionnaire.

Information event card can be considered a mix of diary recording and critical incident technique. As Case (2002:205) pointed out, diary method takes its name from the common type of daily journal in which we record our personal reflections. Diaries are one of the self-report methods of data collection for studying information-seeking behaviour or information needs of users. Nicholas in his book on assessing information needs described diaries as simply self recorded observations - in a methodological sense - and stated that they are

generally used as a substitute for questionnaires and interviews. He mentioned three main advantages for diaries as follow:

- They provide very specific data and very close to the point of action, actions and reactions to events can be recorded at the time of occurrence;
- They are good at getting at people's intentions and then comparing them with the information outcomes;
- Researcher can collect a lot of data over a relatively short period of time by diaries (Nicholas, 2000:131-2).

Barry (1995:120-21) discussed pros and cons of diary as an information collection tool for research on information-seeking behaviour. She maintained that diary is too time consuming, requiring commitment, and not comprehensive. On the other hand it develops self-understanding and brings out tacit knowledge of self as information seeker.

The aim of information-event cards in this research was to gain some information about the actual process of information-seeking behaviour of scientists as they practice it during their working days. Information such as the sort of resources and information they look for, the way they go about the process of information seeking, the incentives or reasons that cause the need for information and drives the information seeking practices would be somehow achievable by means of information-event card.

In this research project a kind of information-event card is used to collect information. Figure 3.8 illustrates the information-event card and it is also reproduced in Appendix 6. Each card was used for recording a single information seeking activity (incident) including the cause of need, the technique(s) used to seek the needed information, the resources used or searched and whether the searcher is satisfied with (found) the results or not. The cards were not used to collect any demographic data but the researcher used an identification code to identify the person who filled the booklet, although the analysis was to be done anonymously.



Figure 3.8: Information-Event Card

**Figure 3.8: Information-Event Card**

The figure displays two sample 'Information-Event Card' forms. The left card is labeled 'Page: 2' and the right card is labeled 'Page: 1'. Both cards have a header with 'Information-Event Card' and a 'Date' field. The left card contains questions 4, 5, 6, and 7. The right card contains questions 1, 2, and 3. Below the right card, it says 'Continues overleaf -->'. The cards are designed to collect data on information-seeking events, including time spent, success, purpose, and search methods.

**Page: 2 Information-Event Card No.**

4. How much time did you spend looking for this information?

5. Did you find the information you have been looking for?

A. Yes, found the information needed.

B. Yes, found some of information needed and will continue search.

C. Yes, found some of information needed and will not continue search.

D. No, no information was found.

E. Other (please explain)

6. For what purpose did you want or need this information?

7. Describe any unusual events in the search, or write any comments you have about this event.

**Page: 1 Information-Event Card Date: / / 2006**

1. Information needed (what subject, in which format):

2. How did you start looking for the needed information?

I. Asking a colleague/friend

If Yes, via  A. Phone  B. E-mail  C. In person

II. Consulting printed resources:

Resources:  A. Book  B. Journal  C. Others, name: \_\_\_\_\_

Location:  A. From library.  B. From your own collection

C. Other, name: \_\_\_\_\_

III. Consulting electronic resources:

A. Electronic journal

B. Other, name: \_\_\_\_\_

IV. General searching on the web

V. Searching a database, please name: \_\_\_\_\_

VI. Other, please specify: \_\_\_\_\_

3. How did you proceed in the process of looking for the information (next steps, if any)?

Continues overleaf -->

After each interview, the researcher introduced the information-event booklet to the interviewees and explained its aim, the way it was supposed to be completed and assured them of the anonymity and the confidentiality of the data. Those interviewees who voluntarily agreed to participate were given a stapled booklet. Each booklet included four blank information-event cards, one completed card as an example, an introduction page explaining the function and the aim of the tool, and the researcher's contact details. The researcher also supplied stamped addressed envelopes to facilitate the return of the booklets. The participants were asked to record the information about their first four information events preferably within the first week after they were given the booklets and post them back to the researcher or contact him to go collect it.

By information event, the researcher meant when one needs some information and look for it, no matter what it is as long as it is related to his/her work as an academician. It could be looking for some papers, browsing a journal, talking to a friend to get some information and so forth. The participants were expected to keep the booklet on their desk and pull it out and fill a card in whenever they fulfilled an information-seeking act.

Totally 45 information-event booklets were handed out to the volunteered participants and after two reminder emails, 24 booklets (53% of booklets) were returned, which included 82 completed information-event cards.

### 3.7.4. Questionnaire

Questionnaire survey is an old, well-established and popular research method in many areas of library and information sciences. Traditionally, pen and paper and mail services (if needed) were used for carrying out questionnaire surveys. Questionnaire as a data collection technique compared to other techniques such as interview has some advantages including: it is cheaper and quicker to administer, it can be done in large quantities, the respondents' answers are not affected by the presence of the researcher (Sudman and Bradburn, 1982), and it is convenient for the respondents (Bryman, 2001). Questionnaire, however has disadvantages too, for example: there is no one available to help respondents if they have difficulty answering questions, there is no opportunity to probe respondents to elaborate an answer, it is difficult to ask complex questions, respondents can easily read all the questions before answering them in order and therefore none of the questions would be independent of the others, and there is risk of missing data due to partially answered questions.

The application of technology in conducting questionnaire surveys probably goes back to early 1990s when computer was applied to help conducting computer-based surveys (Gunter et al., 2002). After expansion of the Internet, researchers began conducting internet-based questionnaires. Internet-based questionnaires, particularly web-based questionnaires, have many advantages, which of course do not come without disadvantages. The main advantages are:

*Low cost:* they are assumed to cost less than traditional print questionnaires. Some past studies confirmed low cost of electronic questionnaire (McCoy & Marks, 2001; Shannon et al., 2002).

*Unlimited geographical coverage:* using post services for distributing print questionnaires restricts its potential geographical coverage, while a web-based questionnaire virtually has no geographical limitations as long as respondents have access to the Internet.

*Speed:* they are faster than print ones due to ease of data collection and processing.

*Quality of responses:* it is more likely to achieve responses of high quality. This is probably because of the interactive nature of the Internet (for example respondents can be provided with help and guidance), and also because respondents associate more anonymity with the Internet and are more willing to reveal their information (Gunter et al., 2002).

*Efficient data processing:* electronic data processing reduces human interferences and hence human mistakes.

Internet-based questionnaires have some disadvantages too. For example the research population is always limited to people who have access to the Internet. Participation in an Internet-based questionnaire survey requires some level of digital literacy.

Sampling is another critical issue in conducting internet-based questionnaires, especially if a probabilistic sample is needed. According to (Couper, 2000), in a probabilistic kind of sampling, researchers choose a sample from a frame population. In probabilistic samples, every member of the frame population has a known, nonzero chance of selection into the survey. Thus while coverage error refers to people missing from the frame (in this case, those without Internet or Web access), sampling error arises during the process of selecting a sample from the frame population, necessitating a means of identifying people on the frame. Therefore, researchers need to know their frame population up to a certain extent and the problem lies in the difficulties of obtaining knowledge about the frame population.

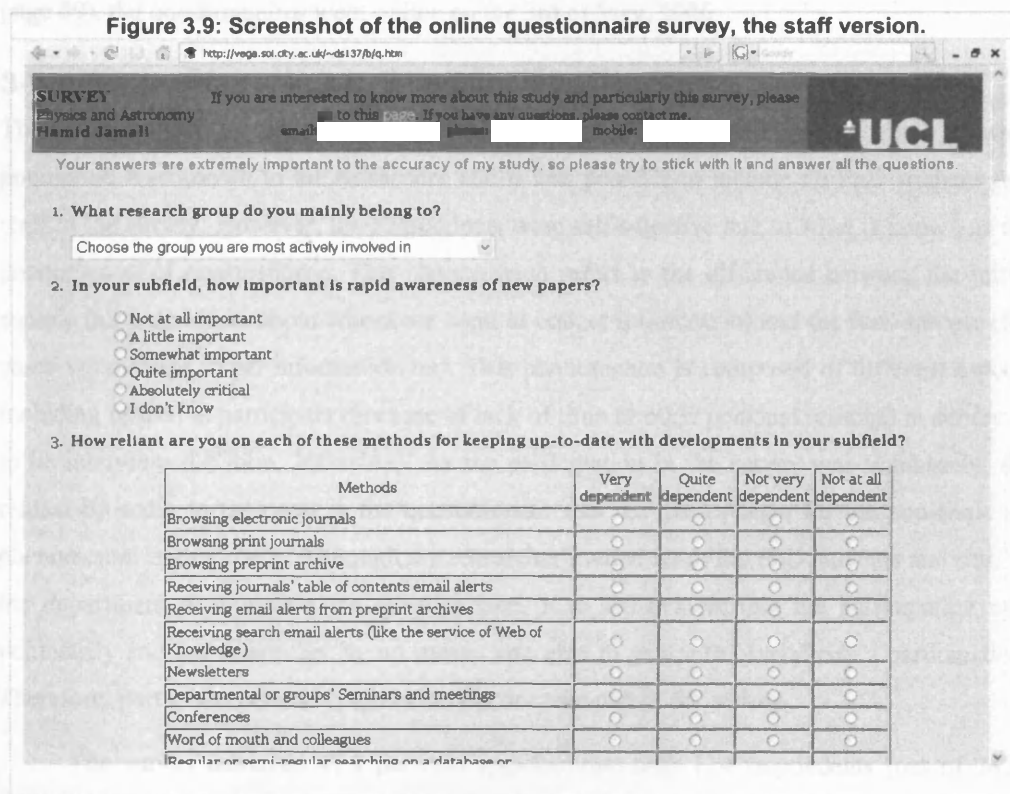
Response rate is thought to be lower in internet-based questionnaires compared to traditional print ones. However, this is arguable. Although some studies showed that response rate in internet-based questionnaires is higher than in print questionnaires (Gunter et al., 2002), some other studies revealed the opposite (Crawford et al., 2002). The truth is that response rate is affected by many factors and it is difficult to issue a general statement about the level of response rate in online or print questionnaire surveys. Among the factors that influence on the rate of response, we can name the topic of research, the length of survey, the characteristics of research population, the incentives of the participants for participation and so on.

Some measures could be taken in order to increase the response rate. Offering financial award is believed to affect the response rate positively. Edwards et al. (2002) showed the odds of response were more than doubled when a monetary incentive was used. The design and the length of the questionnaire have significance in the rate of participation. A user-friendly design would encourage participation. Sending reminders (Kaplowitz, Hadlock, & Levine, 2004) and the right timing for doing so also would increase the response rate (Moss & Hendry, 2002). Assuring respondents of the confidentiality of their responses and their anonymity if appropriate also have positive effect on the rate of participation (Moss & Hendry, 2002).

#### **3.7.4.1. Design of online questionnaires**

After the first phase of the study and preliminary analysis of the data collected through interviewing, a questionnaire survey was conducted to collect some complementary information regarding the issues investigated through the interviews. The questionnaire survey also helped the researcher seek answers to the questions raised from the findings of the interviews.

A self-administered Web-based questionnaire was designed for conducting the survey. Figure 3.9 shows a screenshot of the online questionnaire. The questionnaire had two slightly different versions for students and staff. This slight difference was due to some demographic questions that were different for students and staff. Self-administered means that respondents had to complete the questionnaire by themselves. The choice of web-based questionnaire was due to its advantages in terms of the speed and the accuracy of data collection and analysis as it was already explained. All of the PhD students and members of staff at the department had access to the Internet through their personal desktops.



To increase the response rate and make participation as convenient as possible for the respondents, all of the respondents were given the choice to demand a print version of the questionnaire. Only one respondent (a student) requested a print version, which was delivered to his office with a stamped envelop for its return. A £50 cash prize draw was also offered for increasing the response rate and it was given to the winner (who happened to be a PhD student) in early June after the questionnaire was closed. All respondents were able to decide whether to enter the draw by entering their email addresses. To inform all of the participants

of the results of the draw a short notice was published in the department's weekly email newsletter.

In order to record the responses, a PHP script (PHP: Hypertext Preprocessor) was used to record the responses in a tab-delimited text file, which could be easily transferred to SPSS software (Statistical Package for Social Sciences) for analysis. The usability of the web-based questionnaire was tested with different browsers including Internet Explorer 5 & 6, Mozilla, FireFox, Opera, Safari, and Konqueror. Both the appearance of the questionnaires and the functionality of the PHP script turned out to be compatible with all of the aforementioned web browsers. After piloting the questionnaires with 15 respondents for the reliability issue (see page 89), the questionnaires went online on the 3rd of May, 2006.

#### **3.7.4.2. Sampling and conducting the survey**

There was no need for sampling in the case of the online questionnaire as the whole research population was known to the researcher and it was possible to include all PhD students and staff in the survey. However, the respondents were self-selective due to what is known as the phenomenon of non-response. This phenomenon refers to the difference between the initial sample (all individuals about whom we want to collect information) and the final sample (the cases we manage to get information on). This phenomenon is composed of different aspects including refusal to participate (because of lack of time or other personal reasons) in survey or to be interviewed (Gobo, 2004:441). As the participation in the survey was voluntarily, the refusal by some to take part in the questionnaire was the main reason for the non-response phenomenon in this study. Although the researcher invited all of the PhD students and staff in the department to complete the questionnaire, it is self-evident that the participation was voluntarily and the researcher by no means was able to guarantee everybody's participation. Therefore, part of the research population did not take part in the survey.

The survey achieved 47.1 per cent response rate with 114 respondents (out of 242), which is a good rate by any standard as Hemminger et al. (2007) showed that participation rates range from 3% to 62% for electronic surveys.

To conduct the survey, a personalised email (Appendix 7) was sent to every single respondent with a link to the questionnaire on the 4<sup>th</sup> of May, 2006. The email included some brief information about the purpose of the survey, confidentiality of the data, £50 cash prize, the link to the questionnaire and also a link to a webpage that included some information about the whole research project and the survey particularly for those who wished to know more about the study. The first set of the reminder emails (Appendix 8), which were also

personalised, were sent on the 15<sup>th</sup> of May, 2006, and the second set of the reminders were sent on 22<sup>nd</sup> of May, 2006.

Since one of the main issues in designing questionnaires, especially if a high response rate is to be achieved, is the length of the questionnaire, therefore the researcher tried to keep the questionnaire as short as possible by using mainly closed questions and reducing the number of questions by using two slightly different versions for staff and students. The two version of the questionnaire are reproduced in the Appendix 9 (staff version) and Appendix 10 (PhD student version). Dislike open-ended questions that are not followed by any kind of choice and the answers have to be recorded in full, closed questions offer the respondents a choice of alternative replies. Closed questions are easier and quicker to answer, since no writing is required of the respondent. The respondent simply checks the response that applies to him or her. In addition, closed questions are easy to score and process, when the researcher can easily transfer the results data from the questionnaire to a statistical programme. However, their main disadvantages largely relate to this limited response rate. Nevertheless, closed questions were thought to be appropriate for this research as the survey was conducted after a set of interviews and had a complementary role. The pilot showed that it took one between 8 to 10 minutes to complete the questionnaire.

### **3.7.4.3. The questionnaire's content**

The questionnaire was divided into a few sections (twenty-seven questions in the staff version). The questionnaire started with asking the respondents to mark the research group they belonged to. As one of the main aims of the study was to find out about intradisciplinary differences in information-seeking behaviour, asking this questions at the beginning of the questionnaire was thought to make respondents more conscious about their subfields while completing the rest of the questionnaire. This question was followed by a set of questions covering the following main topics:

#### **3.7.4.3.1. Keeping up-to-date (Q 2-4)**

Respondents were asked to specify how important the rapid awareness of new papers was in their subfields. They were provided with five options, including not at all important, a little important, somewhat important, quite important, absolutely critical, and I don't know. The next question was meant to help find out about the techniques they depended on for keeping up with the developments in their subfields. Eleven main methods were listed, which respondents needed to specify how dependent they were on each (very, quite, not very, not at all). The listed methods were mainly derived from the preliminary results of the interviews. They were: browsing e-journals, browsing print journals, browsing e-print archives, receiving

table of contents (ToC) email alerts, receiving alerts from e-print archives, receiving search email alerts, receiving newsletters, attending departmental meetings and seminars, attending conferences, word of mouth and colleagues, and conducting regular or semi-regular searches. Additionally the respondents were given a choice to mention any other methods not included in the aforementioned list in a textbox. After this question, respondents were invited to rank the top three methods among these methods that they depended on for keeping up-to-date.

#### *3.7.4.3.2. Methods used for identifying research articles (Q 5-6)*

Respondents were asked to tick an option (daily, 2-3 times a week, 1-2 times a month, less than once a month, and never) indicating how often they used each of the eight listed methods for identifying research articles. The eight listed methods were recommendation from friends, ToC email alerts, browsing or searching journal websites, tracking references at the end of papers, searching a general database, searching a subject specific database, searching Google, and searching Google Scholar. Respondents were also invited to name the most used method among these.

#### *3.7.4.3.3. Reading behaviour (Q 7-11)*

To find out any co-relation between information-seeking behaviour and the amount of reading respondents did, respondents were asked to write the approximate number of published articles, and preprints they read in an average month. It was clarified in the question that reading meant going beyond abstract and reading some parts of the article at least. They were also invited to mention if they were personally subscribed to any journal. Those journals and magazines that they might have received through their society memberships were excluded from this question. This question about subscription was removed from the version of the questionnaire that was used for students. The reason was that none of the interviewed students were personally subscribed to any journals and it was reasonable to expect students not to pay for any personal subscription due to high price of scholarly journals. Removing this question helped make the questionnaire shorter for students.

In order to avoid overgeneralisation and try to find more accurate data on their information-seeking behaviour, critical incident technique was used. Respondents were requested to think of the scholarly article they had read most recently and specify how they had found it and how old it was. The question about the methods used for finding the last article read was a semi-closed question in that respondents were able to write their used method in a text box if it was not listed among the eight given options. The options covered different possibilities including: through a colleague, through email alert, through searching

Google, or Google Scholar, or searching a database. One of the options was 'I already knew about it' that implied the article had already been read by respondent and it had been a reread. As to the age of the paper, a range of options from a few weeks to more than 15 years were provided.

#### 3.7.4.3.4. *Statements (Q 12)*

Two statements that were derived from the preliminary results of the interviews were given in the questionnaire and respondents were requested to specify the extent to which they agreed or disagreed with them. The options were strongly disagree, disagree a little, neither agree nor disagree, agree a little, strongly agree, and I don't know. The two statements were:

- I tend to avoid using subject key words and phrases when searching databases to find articles because it brings up too many results.
- If an article is not available online, it's probably not worth the effort to obtain it.

#### 3.7.4.3.5. *Publishing behaviour (Q 13-16)*

Respondents were provided with a textbox to write the number of papers they had published in refereed journals during the past two years prior to the survey. In order to avoid different estimations by respondents they were asked to exclude those papers which were in press at the time of the survey. This question was not included in the students' questionnaire clearly because writing paper is not the main concern for PhD students, especially if they are in the early stages of their studies. The other three questions in the publication section of the questionnaire were related to respondents' interaction with e-print archives. They were asked whether they deposit most of their papers in e-print archives or not, if yes, when and if not, why. Three options were provided with regard to the time of depositing including before submitting paper to journal, at the time of submission, and once the paper is accepted by journal. Regarding the reason for not depositing papers to e-print archives for those who said they did not, four reasons were mentioned in addition to a textbox for those who had another reason other than the given four. The reasons were 'I cannot be bothered', 'I don't see any benefit in it', 'it is not common or a tradition in my subfield', 'the copyright of the journal doesn't allow it'. It should be mentioned that these four options were mainly extracted from the reasons interviewees were given in the interview phase of the research.

#### 3.7.4.3.6. *Characteristics of research subfield (Q 17 & 18)*

It was important to understand the respondents' perception of their subfields. For this end, they were asked two questions. In the first question, which aimed to estimate



interdisciplinarity of their subfield, respondents were asked to specify how often they used the results of research from other disciplines. They had four options that were never, rarely, sometimes, and often. A second question asked respondents what they thought of the journal literature of their subfields. Three main options of this closed question were:

- It's very scattered in many journals and searchable through several databases.
- It's reasonable, not very scattered and not very concentrated.
- It is quite concentrated in a few journals and searchable through a few databases.

'Don't know' option was also available.

#### 3.7.4.3.7. *Problems (Q 19)*

In order to find out about the problems and difficulties that the respondents might have faced in their information seeking activities, five problems were given from which respondents could choose as many as they wished. These were the problems that the interviewees mentioned in the interviews. A textbox also was provided for adding any other problems. Respondents could also choose the option indicating that they did not have any problems. The five mentioned problems were:

- The backfile and older issues of journals are not available online
- There are too many papers out there
- Obtaining papers from obscure journals
- Accessing electronic material from home
- Obtaining conference proceedings
- I don't have any problems

#### 3.7.4.3.8. *Demographics (Q 20-)*

Questions regarding demographic characteristics of respondents were put at the end of the questionnaire because they were easy and fast to answer. Through a number of closed questions respondents were asked to specify their gender and age. These questions were common in both versions of the questionnaire. The staff version of the questionnaire had a few extra questions. The staff members were also required to specify their academic status, and finally the percentage of their time allocated to research, teaching, writing, and other activities.

The participants were also asked to specify the type of their research including theoretical, experimental, observation, instrumentation and a bit of both. A short description of each of these is as follows:

*Theoretical physics*: entails use of mathematical models and abstractions of physics (as opposed to experimental processes) in an attempt to understand nature.

*Experimental Physics*: this part of physics deals with experiments and observations pertaining to natural/physical phenomena.

A bit of both theory and experiment physics:

*Observational Astronomy and astrophysics*: this part of astronomy and astrophysics entails getting data by using different observational devices. It is opposed to theoretical astrophysics.

*Theoretical astrophysics and astronomy*: it is mainly concerned with finding out the measurable implications of physical models. They use a wide variety of tools which include computational simulations and analytical models.

*Instrumentation physics*: this part of physics is concerned with developing instruments for physical measurements.

#### 3.7.4.3.9. Comments

Finally, respondents were given enough space to make any comment regarding their information-seeking and communication activities.

### 3.8. Validity and reliability

Two important issues in any research are validity and reliability of the research methods and data. Validity refers to the question of whether what is measured or recorded in the research is what the researcher intends to measure/record or not. On the other hand, reliability refers to the correctness of measurement/recording and whether it is without error (Tashakkori & Teddlie, 1998:82). These two issues traditionally have their roots in the quantitative or positivist research. Both of these issues have been thought about and taken into account in quantitative and qualitative parts of this research and they are discussed below.

#### 3.8.1. Quantitative part of the research

##### 3.8.1.1. Validity

Validity refers to the issue of whether an indicator (or set of indicators) that is devised to gauge a concept really measures that concept (Bryman, 2001:73). Measurement validity means how much an instrument measures what it is claimed to measure. It is about the accuracy and whether the implementation is correctly indicating what it is supposed to do. There are several

ways of determining if the measurements the researcher uses are valid. Among them the three main ones are content validity, criterion-related validity, and construct validity.

In content validity 'A group of judges (experts) evaluate the degree to which items on a test measure the intended instructional objectives or the content (not useful for constructs that have no specific content)' (Tashakkori & Teddlie, 1998:83). In criterion-related validity, an indicator is compared with another stable measure of the same construct. There are two types of criterion related validity. Concurrent validity is where the criterion variable exists at the present and the predictive validity is where the criterion variable will not exist till later. Construct validity focuses on how well a measure conforms to theoretical expectations (Punch, 2005).

The validity of the questionnaire in this research is mainly based on the content validity. The questionnaire was sent to some experts in the field of information behaviour (such as Professor Carol Tenopir, Professor David Nicholas and Dr. Ian Rowlands) in order to have their judgment on its validity. The researcher received positive feedbacks as well as some suggestions that were implemented in the final version of the questionnaire.

### **3.8.1.2. Reliability**

Reliability basically means consistency and refers to the consistency of measurement. As Hesse noted reliability is the test of measurements consistency over repeated applications. It is another piece of measurement strategy where by the researcher tries to design questions of indicators which represent accurately the concepts. It is the degree to which the results of a measurement accurately present the true magnitude or quality of a construct. There are several ways of determining the extent to which the measurement of attributes are accurate. One main aspect to this consistency is consistency over time (stability). The most obvious way of testing for the stability of a measure is a *test-retest* method. This involves administering a test or measure on one occasion and then re-administering it to the same sample on another occasion. We should expect to find a high correlation between results of the test and retest (Punch, 2005, Bryman, 2001, and Tashakkori & Teddlie, 1998).

Bryman (2001:70) mentions a few problems with this approach to evaluating reliability. For example, respondents' answers in the first administration of the questionnaire may influence how they reply in the second one and this may result in greater consistency between the two administrations of the questionnaire than is in fact the case. Another problem could be that during the time span between the two administrations some factors may change (for example personal financial situation) and therefore the respondents' answer may change. However, Bryman argues that 'there are no obvious solutions to these problems, other than by

introducing a complex research design and so turning the investigation of reliability into a major project in its own right (2001:70).'

To test the stability in this research, the researcher sent 15 questionnaires to 15 friends and colleagues, mainly PhD students. After four weeks, the questionnaire was resent to them. All the responses were entered into SPSS software and the correlation (Spearman's rho correlation test<sup>9</sup>) was conducted and generally the correlation value was more than 0.87, the significance level for all values was at 0.005.

### 3.8.2. Qualitative part of the study

Because of the nature of the data and the fact that there are no measurements in the qualitative research in the sense that there are in quantitative research, validity and reliability in qualitative research are more complicated concepts. Mansourian (2006:74) reviewed qualitative research literature and concluded that: Firstly, there is no agreement on the real meaning of validity in qualitative research literature. Secondly, the meaning and the way of measuring the research validity in qualitative research is different from quantitative research.

As stated before, these two concepts mainly have been developed and employed in quantitative research with the positivist approach and it is difficult to measure them in the same ways in qualitative research. Therefore researchers have paid attention to the need for different concepts in qualitative research, hence the use of terms such as credibility, truthfulness, and dependability. Winter (2000) reviewed the literature for the concepts of validity and reliability and maintained that although researchers realise the need for some kind of qualifying check or measure for their research, they have argued that the term validity is not applicable to qualitative research. Consequently many researchers have espoused their own theories of 'validity' and have often generated or adopted what they consider to be more appropriate terms, such as confirmable, credible, plausible, relevant, representative, trustworthiness or worthy (Hammersley, 1987; Guba & Lincoln, 1989; Mishler, 1990; Wolcott, 1990; Denzin & Lincoln, 1998).

In qualitative research usually the two concepts of validity and reliability are considered in one single but three-dimensional concept of 'credibility'. Credibility has three aspects which are objectivity (neutrality), validity (truthfulness), and reliability (replicability).

#### 3.8.2.1. Credibility

The following three sections explain the measures taken by the researcher in order to assure the three aspects of credibility of the findings of the qualitative part of the research.

---

<sup>9</sup> This test is designed for the use of pairs of ordinal variables (Bryman, 2001:229).

### 3.8.2.1.1. *Objectivity (neutrality)*

Objectivity refers to how unbiased the researcher has been in collecting and exploring the data. The researcher has taken a few different strategies during data collection and analysis to warranty the neutrality of the data and their analysis.

To enhance the objectivity in qualitative research it is recommended to think comparatively and looking for similarities and differences in emerging concepts, collect more evidence, questioning the reality of the collected data, consider the emerging concepts and categories as provisional at the beginning and validate them in the next stages of data collection and keep making comparison (Strauss & Corbin, 1998:43-46). In this research all of the above strategies were employed. Constant comparative analysis for a long time was the main job of the researcher to enhance the objectivity.

This long period of interplay between the researcher and the dataset which was built up gradually and constantly was an important point to ensure the objectivity of the research. Moreover, as mentioned before the whole courses of interview sessions have been tape recorded and then have been transcribed.

The pilot interviews were very helpful to improve the objectivity of the researcher and his ability to interact with participants in a more efficient way. Through analysing the pilot interviews new perspectives appeared in the research and also conducting this stage provided the researcher with new experiences to carry out the main data collection phase with more practical skills. For example, after the pilot interviews the initial hesitation of the researcher about the success of interview as a suitable method for data collection was diminished because he found out that conducting interview and then transcribing is a manageable task for him. The pilot interviews also had some other advantages that warranted the neutrality of the data collection. For example the researcher learned how to avoid steering the interviewees' minds to any direction and just let them talk freely about the main topic. Using open ended questions let participants to freely express themselves and this was another element that helped neutrality of the data.

### 3.8.2.1.2. *Validity (Trustfulness)*

As Lewis & Ritchie (2003:273) stated the validity of findings or data is traditionally understood to refer to the correctness of precision of a research reading. It entails two aspects which are first: whether the researcher is investigating what he or she claims to be investigating (internal validity) and second (external validity): whether the findings are applicable to other groups within the population or to other contexts or settings. Lewis &

Ritchie (2003:275-6) suggested a few methods for validation and verification of the findings, including:

A) Triangulation: it assumes that the use of different sources of information will help both to confirm and to improve the clarity, or precision of a research finding. Triangulation can be of different types including triangulation of sources, which means comparing data from different qualitative methods (e.g. observation and interviews), or methods triangulation, which compares data generated by different methods (e.g. qualitative and quantitative). As it has been explained in this chapter this study is a mixed-methods study. Therefore methods triangulation has been partly used for validation of findings.

B) Member of respondent validation: which involves taking research evidence back to the research participants (or to a group with the same experience or characteristics) to see if the meaning or interpretation assigned is confirmed by those who contributed to it in the first place. In this research once the results of the study were preliminarily written down, the researcher sent the findings to a board of experts composed of five physicists and astronomers in order to verify and validate the findings. The experts had a chance to read the findings of the research and see if there was anything controversial to their knowledge of physicists and astronomers' information-seeking behaviour. The feedbacks received were generally positive and some suggestions and comments were made that were considered in the final writing of the thesis.

### 3.8.2.1.3. *Reliability (replicability)*

Different techniques could be used to enhance the reliability of a research project. Double-coding is one of the suggested methods for assessing the reliability of qualitative research (Miles and Huberman, 1984:60-63). Double-coding means that two researchers code the same dataset separately and then compare the results to find out how similar the codes are. Some formulas have been developed for measuring reliability. Miles and Huberman (1984:60-63), for instance, suggested that: 'Reliability = number of agreements ÷ (total number of agreements + disagreements)'. Although feasible, this could be a problematic method as qualitative research deals mainly with words. Words can have different meanings and imply different concepts. Moreover, two different people might code the same data in two different ways. To reduce the problematic nature of this measurement Pace (2003, 2004) modified the formula by dividing the agreed codes into two groups of 'matches' and 'agreements'. The result was this formula: 'Reliability = (M + A) ÷ (M + A + D)' (Pace, 2003:103). However, this approach was not used in this study. Firstly, because it was not feasible given the limitations that the researcher had in his research. And secondly, the researcher believed that

the quantitative nature of this approach would not be suitable for this study; instead, another method which sounds more appropriate has been used to address the reliability of the study.

There are other different techniques for assuring the reliability of the qualitative research. Lewis and Ritchie (2003) pointed out that there are two levels on which the attention should be paid to ensuring the replicability of the qualitative research. First there is the need to ensure that the research is as robust as it can be by carrying out internal checks on the quality of the data and its interpretation. Second, there is the need to assure the reader/enquirer of the research by providing information about the research process. In this context, questions surrounding the appropriate design and conduct of the research are crucial. Some of these questions are listed by Lewis and Ritchie as below:

- Was the sample design/selection without bias, systematically representative of the target population?
- Was the fieldwork carried out consistently, did it allow respondents sufficient opportunities to cover relevant ground, to portray their experience?
- Was the analysis carried out systematically and comprehensively?
- Is the interpretation well supported by the evidence (2003:272).

Adopting this approach for assuring the reliability of the research, the researcher has well-documented all the process of the research including sampling and analysis in this chapter.

In brief, the researcher believes that the research enjoys a considerable level of reliability, validity and credibility. Apart from the above-mentioned elements, other factors have helped achieve this including: long time reflection on the topic in the first year of the study and before embarking on the main data collection stage, long time data collection without any rush to push the research without having sufficient insight and explanations about the emergent concepts, constant interaction with the data, use of mixed-methods for data collection and triangulation of methods, and consulting the experts in the field on the validity and soundness of the research process.

### **3.9. Limitations of the study**

The study was constrained by some methodological limitations. Generally, due to time and resource constraints of the research project, this research is limited to physicists and astronomers in a particular university. Disciplinary differences exist among scientists in different subjects in terms of information behaviour. The findings of this study can not be entirely generalised to all scientists. Also information behaviour and more particularly

information-seeking behaviour are not simple concepts. They are affected by several factors. For example historical traditions that usually exist in well-established departments such as the UCL department, and the nature and kind of the research carried out in a department might affect information seeking patterns of the scholar. These sorts of organisational cultures and attributes might vary from one department to another. Thus one needs to be cautious in drawing conclusion and generalisations even when the generalisations are restricted only to physicists and astronomers.

There were also limitations in the case of each of the data collection methods that were used. The pool of interviewees could have been better representative of people in different status. Unfortunately it was not possible to have a perfect sample of interviewees from different research groups and with different status (PhD students, researchers, professors and so on). Therefore, one should be cautious in interpreting the finding.

Regarding the information-event cards, the main problem was the lack of commitment for completing the cards and also the fact that participants could have acted very selectively in filling the cards. About 47% of the booklets were never returned and there were some uncompleted cards in those returned. This rate of failure was probably because completing the booklet required time. The participants were not motivated enough to remember to pull the booklet out and fill it out in the occasions when they fulfilled an information seeking action.

In the case of the questionnaire survey, self-selectivity of the respondents might have affected the data. Though a response rate of 47.1 percent is satisfactory for an online survey by any standard, different research groups were not presented in the sample evenly and ideally.

### **3.10. Ethical issues**

All studies involving living human participation must comply with a set of regulations normally determined by their hosting organisations. The University College London has an Ethical Committee, whose approval must be sought for conducting research involving human participation. With regard to the ethical issues, the researcher formally contacted the UCL Research Ethics Committee explaining the research process including its outline, aims and objectives, methodology and role of participants. Due to the anonymity and confidentiality of the data, the research committee officials decided that the current research complies with the requirements of the ethical approval as it can be classified under research type C of the UCL Ethical Committee.

Generally any research involving human participation must have informed consent from participants. Ryen (2004) explained that 'informed consent' means that research subjects



have the right to know that they are being researched, the right to be informed about the nature of the research and the right to withdraw at any time. The other standard ethical issue besides consent is confidentiality. Confidentiality indicates that the researcher is obliged to protect the participants' identity, places, and the location of the research.

The current research complies with all of the aforementioned ethical standards, except the confidentiality of the place and the location that does not apply to the current research. All of the data collected through different methods in this research were anonymous in the first place or were anonymized for the analysis and presentation purposes. The data were kept confidential and in a safe place, with no one having access to them but the researcher. All the participants were aware of the aims and objectives of the research and the kinds of usage that would have been made of the data collected with their participation. They participated in the research (interviews, questionnaire survey, and information booklet) with their own will and had right to withdraw from the cooperation at any stage of their participation without having to have any reason or give any explanation.

In the case of interviews, survey and the information booklets, as stated before, the permission was also obtained from the head of the Department of Physics and Astronomy for approaching the students and staff and seeking their participation.

### **3.11. Summary**

This chapter described the methodology of the research. The study deploys sequential dominant - less dominant mixed methods with the qualitative approach being the dominant approach and quantitative approach being the less dominant one. Four data collection methods are used in the study including interview, online questionnaire survey, information event cards and desk research. The process of data collection, data analysis and sampling methods, and validity and reliability issues were discussed. The chapter ends with a brief discussion on the limitations of the study and the ethical issues.

## CHAPTER 4 - DEMOGRAPHICS OF THE SAMPLE

### 4.1. Introduction

This chapter reports the demographic characteristics of the studied sample including the general demographics of the population of the department, the demographic characteristics of the interviewees, the survey participants and those who participated in the information-event card study. The information about sampling in the case of each of these data collection techniques was presented in the third chapter on Methodology. Here, some more details as tables and figures are presented about the sample.

The data presented here provide some contextual information that helps better understand the results which will be presented in the following chapters. It is appreciated that the information covered in this short chapter could be presented as part of the next chapter (results). However, it was decided that the presentation of the information here as a separate short chapter would make it easier for readers to understand the findings. The fact that the findings include a combination of qualitative and quantitative data is another issue that influenced this decision. After consulting these demographic statistics in a separate chapter, a reader would start reading the results and findings with a clearer picture of the research population in mind.

### 4.2. The case study department

UCL department of Physics and Astronomy had about 129 academic and research staff and about 113 research students at the time of this research's data collection (2005-06). Table 4.1 shows the distribution of the population in the Department of Physics and Astronomy by gender and status. It is worth mentioning that due to the low number of senior lecturers, they were included in the 'lecturer' category. As the table shows the Department of Physics and Astronomy is generally a male-dominant department and this has been reflected in the different parts of the study including the interviews and the survey. Only 25 percent of the population were female and the remaining 75% were male academics and research students.

Table 4.1: Distribution of population by gender and status.

Status	Male	Female	Total	%
Professor	20	4	24	10%
Reader	13	2	15	6%
Lecturer	14	7	21	9%
Senior researcher	10	3	13	5%
Research Fellow	46	10	56	23%
PhD student	78	35	113	47%
Total	181	61	242	100%

### 4.3. Interview participants' profile

The data collection of the interview stage started with the research students; the staff were interviewed in a second phase. Twenty six research students were interviewed in the first phase and thirty members of staff (academics) were interviewed in the second phase. In total 56 participants were interviewed for this qualitative part of the study, which is equal to 23% of the population of the department. This is a reasonably good rate of participation and a suitable sample size. As stated in chapter 3 (Section 3.7.2.1), the size of the population in a qualitative study is not a major concern. However, a larger number of participants in the study might arguably have a positive effect on the validity of the research and can cover different aspects of the phenomenon under study. But considering the time and resource limitations of the study, it was not possible for the researcher to interview more participants.

Tables 4.2 and 4.3 present the list of participants with their profile and their interview sessions. Column with G as header shows the gender, and IL column give the interview length. As it is indicated in this column three of the students were email interviewed. 'Type of Research' presents the type of research that interviewees were involved in including theoretical, experimental or instrumentation physics, a bit of both theory and experimental physics observational astronomy. A description of these categories was presented in Chapter 3. Staff's table has an extra column that gives the status of the interviewees including Professors, Readers, Lecturers, Senior Researchers and Researcher (or research fellows).

Because the data must be presented anonymously, a code is allocated to each interviewee. Students' codes start with an 'S' (for Student) and the first two letters of their research group and a number, so *SAP1* indicates a student in Atmospheric Physics. Staff's codes start with *A* (for Academic staff) followed by the first two letters of their research groups and a number. In the following chapters wherever a quote has been presented from an interviewee, the corresponding code is presented so a reader can match the quote with the

profile of the interviewee. Table 4.2 includes the list of the interviewed research students and Table 4.3 presents the list of the interviewed members of staff.

**Table 4.2: The profile of the participants and the interview sessions – research students.**

<b>No</b>	<b>Code</b>	<b>RG</b>	<b>Type of Research</b>	<b>G</b>	<b>IL</b>
1	SAA1	AA	Theory	M	36
2	SAA2	AA	Theory	M	35
3	SAA3	AA	Theory	F	Email
4	SAA4	AA	Experiment	F	41
5	SAM1	AMOP	Theory	M	45
6	SAM2	AMOP	Experiment	M	40
7	SAM3	AMOP	Theory	M	34
8	SAM4	AMOP	Theory	M	32
9	SAM5	AMOP	Experiment	M	46
10	SAM6	AMOP	Theory	M	37
11	SAM7	AMOP	Theory	F	43
12	SAM8	AMOP	Theory	M	41
13	SAP1	AP	Theory	M	49
14	SCM1	CMMP	Theory	M	37
15	SCM2	CMMP	Experiment	F	39
16	SCM3	CMMP	Theory	F	47
17	SCM4	CMMP	Experiment	M	44
18	SHE1	HEP	Experiment	M	38
19	SHE2	HEP	Experiment	M	39
20	SHE3	HEP	Experiment	M	41
21	SHE4	HEP	Experiment	M	42
22	SHE5	HEP	Theory	M	Email
23	SHE6	HEP	Experiment	M	Email
24	SOS1	OSL	Instrumentation	M	50
25	STM1	TMP	Theory	F	37
26	STM2	TMP	Theory	M	35

Table 4.3: The profile of the participants and the interview sessions – academic staff.

No	Code	RG	Status	Type of Research	G	IL
1	AAA1	AA	Lecturer	Observation	F	40
2	AAA2	AA	Lecturer	Theory	M	42
3	AAA3	AA	Professor	Both	M	56
4	AAA4	AA	Professor	Theory	M	63
5	AAA5	AA	Professor	Observation	M	54
6	AAA6	AA	Reader	Observation	F	49
7	AAA7	AA	Researcher	Theory	F	42
8	AAA8	AA	S. Research	Observation	M	53
9	AAM1	AMOP	Researcher	Theory	M	47
10	AAP1	AP	Professor	Theory	M	51
11	AAP2	AP	Researcher	Experiment	M	44
12	ACM1	CMMP	Professor	Experiment	M	71
13	ACM2	CMMP	Professor	Theory	M	41
14	ACM3	CMMP	Professor	Theory	M	65
15	ACM4	CMMP	Reader	Theory	M	50
16	ACM5	CMMP	Researcher	Experiment	M	52
17	ACM6	CMMP	Researcher	Experiment	M	39
18	ACM7	CMMP	Researcher	Theory	M	41
19	ACM8	CMMP	Researcher	Theory	M	42
20	ACM9	CMMP	Researcher	Theory	M	39
21	ACM10	CMMP	S. Researcher	Theory	M	46
22	AHE1	HEP	Researcher	Experiment	F	48
23	AHE2	HEP	Lecturer	Experiment	M	43
24	AHE3	HEP	Professor	Theory	M	58
25	AHE4	HEP	Reader	Experiment	M	52
26	AHE5	HEP	Researcher	Experiment	M	54
27	AOS1	OSL	Researcher	Instrumentation	F	49
28	ATM1	TMP	Researcher	Theory	M	55
29	ATM2	TMP	Researcher	Theory	M	52
30	ATM3	TMP	Researcher	Theory	M	50

Forty-seven percent of the participants were research students and 53% were academic staff. Twenty-five percent of the participants were Researcher, 14% were Professors, and 4% were Lecturer. Readers and Senior Researchers each accounted for 5% of the interviewees (Figure 4.1). 19.6% of the interviewees were male and 80.4% were female.

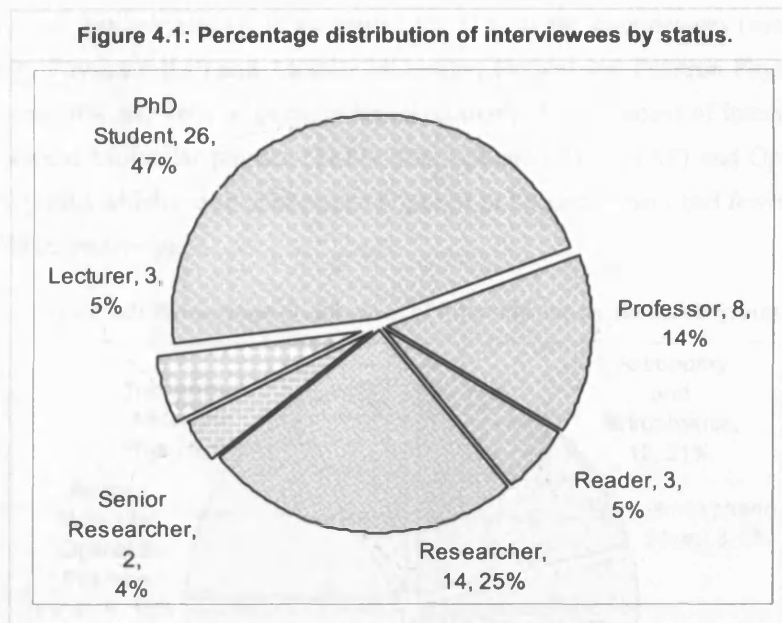
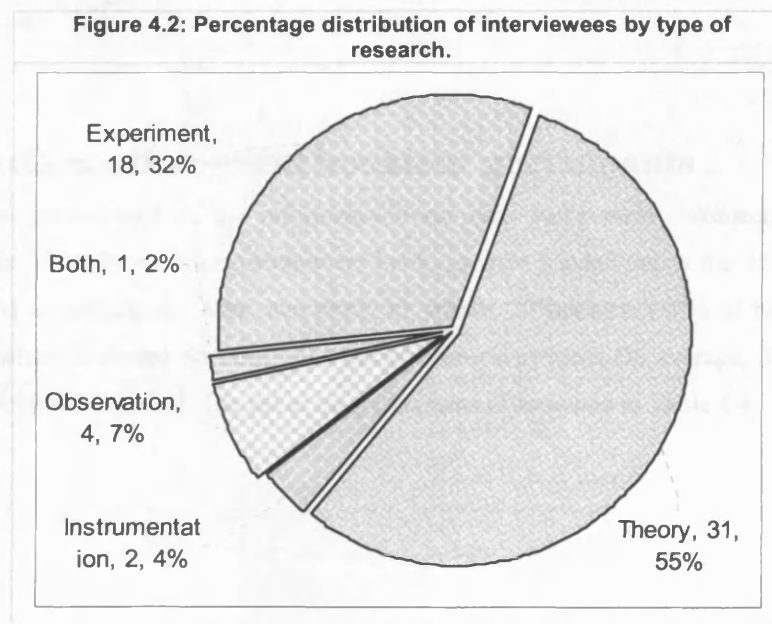


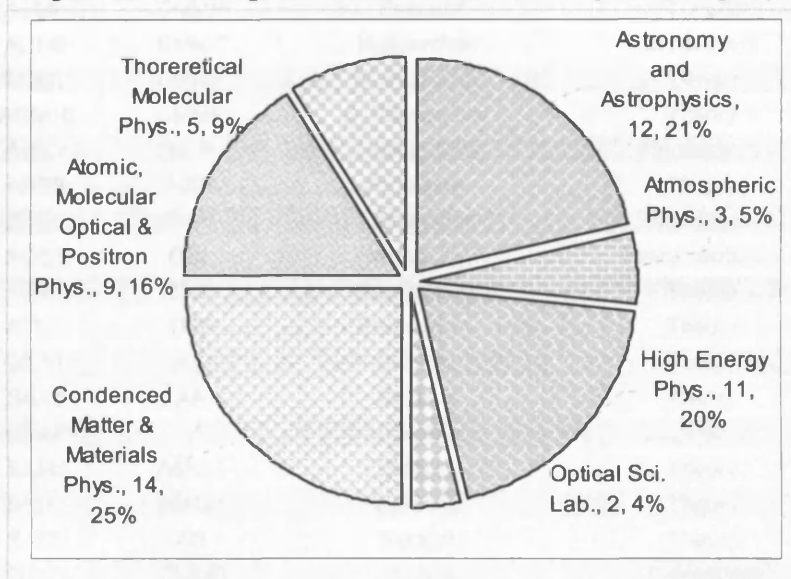
Figure 4.2 shows that 55% of the interviewees were theoretical physicists, 32% experimentalist, 7% were involved in observational astronomy, and 2% stated that their work entailed a bit of both theory and experiment and 4% were instrumentation physicists.



The interviewees belonged to different research groups in the department. Twenty-five percent of them were members of Condensed Matter and Materials Physics (CMMP).

Astronomy and Astrophysics (AA) accounted for 21% of the interviewees (see Figure 4.3). High Energy Physics (HEP) and Atomic, Molecular, Optical and Positron Physics (AMOP) accounted for 20% and 16% of participants respectively. Nine percent of interviewees were from Theoretical Molecular physics (TMP). Atmospheric Physics (AP) and Optical Science Laboratory (OSL) which are smaller research groups in the department had fewer participants (5% and 4% respectively).

Figure 4.3: Percentage distribution of interviewees by research group.



#### 4.4. Information-event booklets' participants

Those who participated in the information-event card study were volunteers among the interviewees. Totally 45 information-event booklets were handed out to the 45 interviewees who agreed to participate. After two reminder emails, 27 booklets (60% of booklets) were returned, which included 88 completed information-event cards (in average, 3.2 completed cards per returned booklet). The list of the participants is presented in Table 4.4.

Table 4.4: The profile of the participants who completed information-event cards.

No	Code	RG	Status	Type of Research	G
1	AAA2	AA	Lecturer	Theory	M
2	AAA5	AA	Professor	Observation	M
3	AAA6	AA	Reader	Observation	F
4	AAM1	AMOP	Researcher	Theory	M
5	AAP1	AP	Professor	Theory	M
6	ACM4	CMMP	Reader	Theory	M
7	ACM5	CMMP	Researcher	Experiment	M
8	ACM7	CMMP	Researcher	Theory	M
9	ACM10	CMMP	S. Researcher	Theory	M
10	AHE1	HEP	Researcher	Experiment	F
11	AHE3	HEP	Professor	Theory	M
12	AHE5	HEP	Researcher	Experiment	M
13	AOS1	OSL	Researcher	Instrumentation	F
14	ATM1	TMP	Researcher	Theory	M
15	ATM3	TMP	Researcher	Theory	M
16	SAA1	AA	Student	Theory	M
17	SAA3	AA	Student	Theory	F
18	SAA4	AA	Student	Experiment	F
19	SAM6	AMOP	Student	Theory	M
20	SAM7	AMOP	Student	Theory	F
21	SAP1	AP	Student	Theory	M
22	SCM2	CMMP	Student	Experiment	F
23	SCM4	CMMP	Student	Experiment	M
24	SHE1	HEP	Student	Experiment	M
25	SHE3	HEP	Student	Experiment	M
26	SHE5	HEP	Student	Theory	M
27	STM2	TMP	Student	Theory	M

Table 4.5 shows the number of booklets (participants) and cards by status and gender. As we can see the majority (74% or 20 out of 27) of the participants were male and only 26% (7 out of 27) were female. The largest group of female respondents were also PhD students. However, this is expected as it matches (approximately) the proportions of the interview sample as well as the general population of the Department of which 75% were male.



**Table 4.5: Number of completed information-event cards by participants' status and gender.**

Status	Gender	No. of booklets	No. of cards
PhD student	M	9	29
	F	3	10
Research Fellow	M	4	13
	F	1	3
Senior Researcher	M	2	7
	F	-	-
Lecturer	M	3	9
	F	1	4
Reader	M	2	6
	F	-	-
Professor	M	2	7
	F	-	-
Total		27	88

#### 4.5. Survey participants

Table 4.6 shows the size of the research population and the number/percentage of respondents. All of the research students and the academic staff in the department were invited to take part in the survey. However, only 114 respondents completed the online questionnaire which is equal to a 47.1 percent response rate. This was considered good given that the academic Web-based surveys' participation rates range from 3% to 62% for electronic surveys (Hemminger et al., 2007). 57% of respondents were research students and 43% were members of staff.

**Table 4.6: The size of the sample and the response rate of the survey.**

	Total	Respondents	
	Population	No	%
Staff	129	49	43
PhD Students	113	65	57
<b>Total</b>	<b>242</b>	<b>114</b>	<b>47.1</b>

As we can see in Table 4.7 more than 70 percent of respondents were male. This is, as stated before, a reflection of the general population of the department, of which 75% were male academics and research students.

**Table 4.7: Distribution of the respondents by gender.**

Gender	No	%
Male	81	71.05
Female	33	28.95
<b>Total</b>	<b>114</b>	<b>100</b>

Table 4.8 shows the distribution of respondents by age. More than two-third of them were 34 and fewer years old. This is because a high percentage of respondents as we can see in Table 4.9 were research students (57%) and also research fellow (17%). There were also ten professors among respondents. Unfortunately, the researcher does not have the demographic data about age of the whole population of the department and cannot weight the data in Table 4.9 against the whole population. However, as Table 4.1 shows 70% of the population of the department are PhD students or research fellows which are quite reasonably expected to be 34 or less. This explains the high rate of respondents in the age category of '34 and under'.

**Table 4.8: Distribution of the respondents by age.**

Age	No	%
34 and under	88	77.2
35-39	5	4.4
40-49	10	8.8
50-59	7	6.1
60 and more	4	3.5
Total	114	100

**Table 4.9: Distribution of the respondents by status.**

Status	No	%
PhD Student	65	57
Research Fellow	20	17.5
Senior Researcher	6	5.3
Lecturer	11	9.6
Reader	2	1.8
Professor	10	8.8
Total	114	100

Table 4.10 shows the distribution of respondents by the characteristic of their research. The highest number of respondents belonged to those involved in theoretical research in the field of physics (31.6%), followed by 33 respondents (28.9%) who did experimental research in physics. The smallest proportion belonged to those who did instrumentation research with only just 4 respondents fell into this category. Seven main subfields of physics and astronomy (research group entities inside the department) were used to categorise the respondents (Table 4.11). Condensed Matter and Materials Physics (CMMP) accounted for 31.6% of the respondents. CMMP is the biggest research group in the department and it encompasses a considerable number of smaller research groups that research on very specific topics. After CMMP, Astronomy and Astrophysics accounted for the second highest number of respondents with 22 (19.3%) respondents. This research group also covers many smaller research groups such as hot stars, star formation and so on. The smallest number of

respondents belonged to the Optical Science Laboratory with 3 respondents who all do instrumentation-kind of research.

**Table 4.10: Distribution of the respondents by type of research.**

Type of Research	No	%
Theory (physics)	36	31.6%
Experiment (physics)	33	28.9%
Observation (astrophysics & astronomy)	13	11.4%
A bit of both	10	8.8%
Theory (astrophysics & astronomy)	18	15.8%
<b>Instrumentation</b>	<b>4</b>	<b>3.5%</b>

**Table 4.11: Distribution of the respondents by research group.**

Research Group	No	%
Atmospheric Physics (AP)	11	9.6
High Energy Physics (HEP)	18	15.8
Condensed Matter and Materials Physics (CMMP)	36	31.6
Astronomy and Astrophysics (AA)	22	19.3
Theoretical Molecular Physics (TMP)	11	9.6
Atomic, Molecular, Optical and Positron Physics (AMOP)	13	11.4
Optical Science Laboratory (OSL.)	3	2.6
<b>Total</b>	<b>114</b>	<b>100</b>

## 4.6. Summary

This chapter demonstrated an overall picture of the main features of the participants in this study. The study has used three main data collection techniques including interview, survey and information-even cards. The Department of Physics and Astronomy at UCL is a male-dominant research oriented department. The department had 113 research students and 129 academic staff at the time of data collection. Twenty-three percent of the population participated in the interviews including 26 research students and 30 members of staff. Eighty eight information-event cards were completed by 11 students and 16 members of staff. And the survey achieved a 47.1 percent response rate as 114 people participated in the online questionnaire survey.

The following chapter, the Results, will present the findings of the study.

## CHAPTER 5 - RESULTS

### 5.1. Introduction

In this chapter, the results of 56 semi-structured face-to-face interviews, questionnaire survey (114 respondents) and information-event card study (27 participants, 88 cards) are presented. The findings are presented under thematic subheadings which are initiating research, methods used for identifying articles, keeping up-to-date, problem-specific information-seeking, accessing information, changes in information-seeking behaviour, reading behaviour, publishing behaviour, and problems and difficulties. Under each heading, all the relevant results that have been obtained using different methods are presented.

There are a few issues to be addressed in the introduction before moving on to the results. First is the logic behind the presentation of the results in this chapter. Second is the structure of the chapter and the source of the headings under which the results are presented. And finally how different types of data obtained through different methods have been treated for the final presentation and how different qualitative and quantitative datasets are related to one another. These issues are discussed below under the following three subheadings.

#### 5.1.1. Presentation

Probably the most challenging issue in this study was to find a suitable way for presenting and structuring the results. The cause of this challenge was the fact that the study uses mixed-methods for data collection and a mixed set of data, rich but of different natures, was collected. To weave a combination of qualitative and quantitative data and present in a coherent way is a challenging task.

There were a few logical possibilities for the presentation of the results in this study. The most logical option, which has been applied in the study, was to form the results chapter around the research questions and objectives and the issues that have been investigated. This means presenting all of the findings by all methods on each area together. The researcher decided that this was the most suitable approach i.e. to present the findings thematically. It is appreciated that combining qualitative data and quantitative data can be problematic and complicated; however, it has the advantage of creating a more comprehensive picture of the findings as they complement one another.

There were two other options that were ruled out by the researcher. The easiest option was to structure the results based on the methods used. In this case the findings of the interviews and the survey could have been presented completely separately. But to form the results based on the various methods used would not have produced the best outcome because the study is not a methodologically-driven one. Moreover, separating the findings of the different methods, while they discuss the same issues, would have made it difficult for readers to understand and follow the findings. This way of presentation could have made it hard for readers to see the whole picture and they might not have seen the wood for the trees.

The other option was to present all the findings about each research group together. Therefore, for example all the findings about high energy physicists including the findings of the interviews and the survey would have been presented together. This could have been a sensible way of presenting the results as one of the aims of the study is to do intradisciplinary comparisons. However, this could have been problematic due to the nature of the quantitative data. While it would have been easy to divide the results of the interviews based on the research group of the participants, dividing the results of the survey would have led to the presentation of duplicated data under each research group. For example, one figure or diagram might have had to be discussed under all of the research groups. Moreover, generating a cluster analysis (see Appendix 11) using the survey data showed that there were not any clear clusters in which users from a single or similar subject areas dominated. The other defect of this method was that intradisciplinary comparison is just one of the main goals of the study and presenting the results based on research groups would have overshadowed the other aspects of the study. However, it should be mentioned that to soundly achieve this goal the research groups will be profiled in the conclusion chapter and the findings about their information-seeking behaviour will be presented there.

### 5.1.2. Structure of the results

As the thematic approach was adopted, the chapter is composed of some major subheadings under which the findings obtained through different methods are presented. The headings are based on the main issues investigated through the interviews and the survey, which in return were informed by the objectives. Having consulted the existing literature to find a pattern for presenting the results, the researcher realised that the literature does not suitably provide that. The journal articles that present the findings of *quantitative* data are rich in this regard. This is because in quantitative studies the statistical results are normally presented as diagrams and tables organised around the questions asked in a survey (see for example Rowlands and Nicholas, 2005). On the other hand, journal articles that present *qualitative* studies tend to bypass the results and present the section which is the final interpretation of the results and

conclusions. Therefore, it is hard to find a qualitative paper that is organised thematically. They tend to present the final theory or model or the interpretations that have been drawn from the findings. For instance what is presented in Ellis (1989) or Ellis, Cox and Hall (1993) is the model derived from the findings. This is what the Conclusions chapter of this study is meant to fulfil.

As the focus of this study is mainly research-related information-seeking behaviour of physicists and astronomers and certain aspects of their scholarly communication in the context of information behaviour, this focus on research drives the structure of the results. Conducting research entails different information seeking activities that can be categorised into three groups. To initiate a new research project one needs naturally to familiarise oneself with the specific subject at hand and this requires going back to the literature in order to obtain the background knowledge and to set the scene for the study. During the course of a research, researchers need to maintain their knowledge up-to-date with regard to the developments of the area. Also from time to time one may face an information gap that can be filled by seeking some information, whether it is a fact finding, or solving a specific problem or obtaining some general background information for which a need occurs during the course of the research. Close to the end of the research and when it comes to preparing the final output of the research which normally materialises in the form of scholarly publications (mainly journal articles) again some information needs emerge, hence some information seeking activities. Along side these information seeking activities, there are other important issues such as accessing information (which can be considered as part of the broader area of information-seeking) and reading behaviour through which users use the information that have sought and found. Publishing behaviour has also been included in the study as the focus is on research and publishing is when the end-product of the research is produced.

Following this logic, the chapter starts with a section on initiating research in which the information seeking activities related to the start of research are discussed, including seeking literature at the beginning of a piece of research. Methods used for identifying journal articles are discussed afterwards as this is related to the use of literature. The next section discusses the methods used for keeping up-to-date. Then the other group of information seeking activities that deal with problem-specific information seeking are described. Two related issues to information-seeking behaviour, which are accessing information and changes in information-seeking over time, are discussed next. After that, the reading behaviour section sheds some light on the way physicists and astronomers use scholarly papers. Publishing behaviour is an important part of the scholarly communication of physicists and astronomers and it represents the last stage of a piece of research which is producing the research output.

Publishing behaviour is discussed afterwards. Problems and difficulties that users face in their information seeking activities are the issues discussed last. The chapter concludes with a summary.

### 5.1.3. Interweaving the data

Having adopted the discussed approach for the presentation of the results, another challenge arose. This challenge was because of the relation between different types of data. In some cases such as methods used for keeping up-to-date and problems in information-seeking, the quantitative data triangulate the qualitative data; in some cases such as reading behaviour the quantitative data are complementary to the qualitative data in that qualitative data show how physicists read while the quantitative data provides some statistics on the amount of reading and its variations among different participant groups. There are also issues that were investigated merely by the qualitative data (such as role of conferences and colleagues) or the quantitative data (such as methods used for identifying articles in connection to the age of articles).

Of course this has been an intentional decision by the researcher so that overall the dataset presents a picture of information-seeking behaviour of physicists and astronomers as comprehensive and accurate as possible within the limitations of the study. The reason for investigating some issues merely through qualitative or through quantitative data has been partly due to the limitations of the study. For example including all of the issues in the interviews would have made the interviews too long and would have reduced the participation. The same is true about the questionnaire survey, too long a survey would have made participants reluctant to participate in the study. The other reason was the nature of the issues investigated. For example the quantitative approach is more suitable for investigating methods used for identifying articles as it objectively provides statistics and the possibility for investigating the correlation between methods used and the other factors such as the age of articles. Using a quantitative approach for investigating this issue is well-established as well (see for example several studies by Carol Tenopir). On the other hand, issues such as the role of conferences and colleagues in the information-seeking behaviour of the participants is more suitable to be researched through qualitative methods as participants get a chance to explain and delineate the issues; and the interaction between the interviewer and the interviewees would result in a more detailed account of the issue compared to what is achievable through simply ticking a few preset options in a questionnaire.

In presenting the data, wherever it was possible to interweave the qualitative and quantitative data, this has been done. For example the qualitative and quantitative data about

keeping up-to-date complement each other and they have been presented in a logical order in a way to make it easy for readers to understand.

However, while reading the results in this chapter, users may find themselves switching from qualitative data to quantitative and vice versa. It is appreciated that this may cause a cognitive barrier in understanding the data and might make it hard for readers to follow the storyline. Measures have been taken in order to minimise this effect. Although quantifying the qualitative data is not a wise thing to do in most cases, it is rational to add some figures in order to give a sense of proportion to readers. This has been done wherever appropriate in the qualitative results so the qualitative and quantitative results will look slightly less heterogeneous. Also, to avoid causing confusion as to what data source has been used in each section of the findings that are presented in this chapter, Table 5.1 lists different issues investigated and their data source. The table is meant to clarify the data source and serves as a point of reference for the readers so one can easily find out what dataset was used for each issue or information seeking characteristics discussed in the results.

It must also be noted that the data presentation is somewhat driven by the findings themselves. Wherever clear differences have emerged in the information-seeking behaviour of the participants by different status or different research groups, they were discussed separately so that a clearer picture of their information-seeking behaviour is presented to the readers. For example, in the section on initiating research staff and students are discussed separately because clear differences emerged between them in this regard.



Table 5.1: Methods used for investigating different issues.

Characteristics of information behaviour investigated	Interviews	Survey	IEC
Initiating research	✓		
Methods used for identifying articles	✓	✓	
By frequency of use		✓	
By users' academic status		✓	
By users' gender		✓	
By users' type of research		✓	
By interdisciplinarity of the field		✓	
By scatter of the field's literature		✓	
By age of article		✓	
By users' research group		✓	
Keeping up-to-date	✓	✓	
Methods used for keeping up-to-date	✓	✓	
By dependency on various methods		✓	
By importance of keeping up-to-date		✓	
By age of participants		✓	
By academic status of participants		✓	
By type of research		✓	
By interdisciplinarity of the field		✓	
By scatter of the literature		✓	
By research group	✓	✓	
Specific groups (AP, HEP...)	✓		
Role of conferences	✓		
Role of colleagues and research environment	✓		
Problem-specific information seeking	✓	✓	✓
Web searching	✓	✓	✓
Accessing information	✓		
Changes in information seeking over time	✓		
Reading behaviour	✓	✓	
Reading quantity		✓	
Reading and other activities		✓	
By research group	✓	✓	
By methods used for identifying articles	✓	✓	
By importance of keeping up-to-date	✓	✓	
By age of the material read	✓	✓	
Reading from screen	✓		
Role of abstracts	✓		
Validity of information	✓		
Publishing behaviour	✓	✓	
Writing approach	✓		
Publishing approach, in journals	✓		
Publishing approach, in e-print archives	✓	✓	
Problems and difficulties	✓	✓	✓

## 5.2. Initiating research and seeking information

To start a new research project one needs to go through an initial familiarisation process and acquire some information about the subject and the area of the research. This process is different for staff and students. This is because of the nature of the research students' involvement in the research process. As will be explained below, research students tend to join bigger research projects in the department and they are guided by their supervisors and colleagues in their research. This makes the nature of their research process different from the staff's. Therefore, the staff and the students are discussed separately below. The data is based on the answer of the interviewees to a question as to how they go about starting a new research project.

### 5.2.1. Staff

As expressed by about 85 percent of the staff interviewed, there are rarely big changes in the research interests of physicists and astronomers. Physicists and astronomers like other scientists are specialists; however, scientists may shift their research interests due to the scientific trends that can be influenced by new discoveries, changes in research funding policies and other factors. Research is a long and continuous process. New research projects tend to emerge from the outcome of the previous projects and they come up while scientists are working on their current research.

I mean you don't just suddenly start a new research project. Usually somebody carries on through something you had been doing already, so usually it is something that comes up while you are working on something else, or you have a new piece of data or something. It is not something you go actively looking for. It is something that happens; it comes about while you are looking at something else usually. [AAP1<sup>10</sup>]

However, there are minor variations among the subfields of physics and astronomy in terms of the way they come up with new research projects. In astronomy, researchers may find themselves interested in the relevant research questions that are raised in review articles and also during friendly conversations with colleagues or in conferences. They pursue these questions and they might lead to new research projects. In observational astronomy, the ones who build the research teams have a problem and a proposed solution for which they get funding and the team work on that problem [AAA3].

---

<sup>10</sup> These codes refer to the interviewees. Codes start with 'A' refer to staff and the ones starts with 'S' refer to PhD students. The next two letters refer to the research group of interviewees, e.g. AAA is a member of staff in Astronomy and Astrophysics group. A list of interviewees and their correspondent codes was presented in the 4<sup>th</sup> Chapter to which the reader can refer for details.

Research in theoretical high energy physics is very ad hoc. Theoretical high energy physicists tend to work in small groups and new ideas for new research projects may come out of their chats. While experimental high energy physicists tend to work in large groups and run long projects that can last for 15 years. The way new research is started in experimental High Energy Physics (HEP) is that during the research group meetings, some present the findings of new theoretical papers in HEP. Theoreticians suggest solutions for existing problems and experimentalists investigate whether what theoreticians propose is feasible and doable experimentally or not. To decide to build an experiment is a long and thorough process as experiments in HEP are very expensive and could cost billions of dollars, involving researchers from several countries.

What tends to happen with new research projects is there is... there tends to be a series of meetings where people will present their ideas in power point. And there tends to be several of those and the precursor to that is often that there will be some theoretical papers with the motivation for a new project or the potential for this new project. And then as experimentalists we tend to get together and say is it feasible, can we build it, how much will it cost, what are the time lines going to be, can we actually achieve what is worthwhile you know to prove or disprove the theoretical motivation. [AHE4]

The information-seeking behaviour of physicists and astronomers in the early stages of a new project varies somewhat based on the subfield of physics and astronomy in which they work. This initial familiarisation with the subject area of a new project may be done in a few different ways. It is reasonable and sensible to expect someone who is about to start a new research project to review the literature at the beginning of a project in order to make sure that enough background knowledge has been obtained and that what he or she is doing is right. One would want to make sure that the research is based on past experience and it follows a sound line of research. However, the way different people go about this task is different.

To find out more about literature searching the interviewees were asked 'at what stage(s) of a research project they search for literature, why and how?' The interviews showed that physicists and astronomers mainly seek literature two times during a research project, but with different aims; and therefore the nature of their literature search is different: one major search at the beginning, another search for the literature close to the end or at the time of writing (some also do searches for literature during the course of the project). Here the initial search is explained, the search at the time of writing will be explained in the publishing section (page 179) and the occasional searches during the project will be discussed in the problem-specific information seeking section (page 154).

The interviews showed the main and the most comprehensive search is conducted at the outset of a research project. Search here is a generic term and does not merely indicate an

actual search in a database, as some people may track references to seed papers for finding relevant literature. The aim of the comprehensive literature search at the outset of forming a research project is to gain as much information as possible on the subject, to avoid repetitive work and mistakes. To start a research project in physics and astronomy, researchers naturally need to write a grant proposal and get funding. In order to write a grant proposal, a thorough literature search must be carried out. The following quotation explains why this is. The interviewee maintained that grant applications in these days are over-subscribed and only one application out of ten is likely to succeed so it is a very, very competitive business in getting money.

You will only get that grant application approved if you can show that you have a good track record in the area, you understand the problem, you know the literature and you know you identified a particular problem and you have identified a way of trying to solve it, so in writing grant applications literature knowledge and the area knowledge is an integral part of that particular exercise. [AAA3]

However, it must be said that the methods used for approaching the literature at the beginning of a new research project are not the same for all researchers and for all research groups and there are variations. In the astronomy section of the department, there are more interpersonal relations compared to the physics section and it appeared that people rely on each other's information a great deal. A few of the astronomers interviewed stated that the first thing they would do at the beginning of forming a new research project is to approach their colleagues and find out if they know anything about the subject, whether they have worked on that subject or they know people who have worked or key papers on the subject. This might guide them towards key resources such as good review papers and names of some important researchers in that area. If some key papers are identified then chasing references of those papers for finding, or findings citing papers that have cited those papers would be the second step. And if the result of chatting with colleagues is identifying some authors in the field or nothing then the next step would be conducting searches in Astrophysics Data System (ADS) database. The ADS is a very comprehensive database in the astrophysics and related subjects. It contains full-text of most of the astronomical literature. However, as a matter of personal preference some astronomers may chat with colleagues in parallel with using ADS (for example the interviewee AAA1) and some may chat with colleagues first and then start their search in ADS (for example the interviewee AAA2).

So probably the first thing I would do is search for literature on the related subjects. So the first thing I would do is go through ADS. But probably the second thing I would do, maybe in parallel, is chat with the people in the corridors or maybe from different groups. We often have here social meetings, coffee meetings, for example a drink Thursdays at five where you end up meeting people and you can chat about new ideas. I think that's how we start. [AAA1]

Experimentalist High Energy Physicists tend to work in large groups that may include up to hundreds of researchers. In the early stages of a project a few of researchers do the literature review and present summaries to the other researchers. There are constant meetings among people who are involved in a project and everybody shares the summaries of their findings, whether they are on the technical devices or the theoretical aspects of the experiment. Spires<sup>11</sup> and the experimental high energy physics section of arXiv known as hep-ex<sup>12</sup> seem to be the two main sources for finding papers in experimental high energy physics.

In theoretical high energy physics, things are a little different. The groups are smaller. Although senior researchers who are in charge of a research study may delegate parts of the information-seeking tasks that must be done in early stages of a study, sometimes even professors find it helpful to do some of the information-seeking tasks. This perhaps is a reflection of the nature of the theoretical research which requires the researchers to be well-informed of the background of the problem they want to work out. The following quotation is from a theoretician professor in high energy physics who explains why even professors sometimes need to do their own literature search in order to get more acquainted with the literature.

In all research you have to acquaint yourself with literature. That's vital because you may think you know the literature, but you need to do some sort of literature search. Very often if you have a research student you might get them to do that. But it's always good to do things like that yourself as well. So that you should be pretty familiar with what's gone on before, perhaps approaches other people have tried, may not have been quite successful etc. Then you are ready to start. You've got a better feeling of what the real problem is you want to solve. It's essential to do some background reading. [AHE3]

Theoreticians still rely mainly on reference tracking to dig up the background of a theory or model or a specific problem in physics. However, chasing citations for finding reference material means that researchers need some recent key papers to start with. For finding those key papers (most probably review articles) they would search databases, especially the high energy section of arXiv.

People in the Optical Laboratory rely on subject searching in general databases to find background information and scientific papers on their projects. Interestingly they tend to rely somewhat on Google for this purpose too. This is because their research has several aspects including astronomy (as optical devices are made for telescopes), physics, engineering and some technical aspects. As it is discussed in the Problems section (see page 197), finding technical information especially if they are held by commercial companies could be difficult;

---

<sup>11</sup> <http://www.slac.stanford.edu/spires/>

<sup>12</sup> <http://arxiv.org/archive/hep-ex>

and UCL library also does not subscribe to all of the relevant journals that these groups of physicists wish to have access to. Therefore, they resort to Google knowing that it would enable them to find the information whether they are in scientific papers or on an author's website. A member of staff in the Optical Laboratory described why in their field Google is considered a good tool for finding information. It is because the scientific information, even papers, are stored on places such as personal web pages and open access repositories as well as journals to which a user may not have access to.

What I do like to use if I'm looking for more the astronomy side I use ADS which is the NASA database. And then sometimes I just search on Google. I just noticed this, well especially in optics, people tend to, because some of the journals you have to pay for, there's a subscription and you can't just download them for free, but you normally have the personal websites, so a lot of people will put the paper on their website, say. Sometimes that's the best way of finding the paper if you just type in... especially if you know an author or a group that are doing a particular research, you sometimes just go to their website and just see their recent papers. [AOS1]

There were also variations in the information-seeking behaviour of members of staff with different status in the initial familiarisation stage. The interviews showed that senior members of staff such as professors and readers and those who are older with more experience and knowledge tend to rely less on literature searching. They tend to be more focused in their information-seeking behaviour and rely on methods such as personal communication with colleagues and chasing references rather than conducting a search in a database. When was asked about his looking for the literature in the early stage of a project, a professor stated that he had not done a literature search for thirty years:

...you build up that information through working in the field. I probably haven't done a literature search as such for 30 years, because when you're working deeply within a field then you tend to know the people who are working in that area and you tend to hear what is actually happening through the people you're collaborating with, through going to conferences, through reading the literature when you're sent papers. So I haven't done a real literature search for a long time and so, as I say I may be untypical. [ACM1]

Some other professors made similar remarks by confirming that they do not tend to do literature search and prefer to get their students to do it.

Yeah, it's not a search; a search is a bit of an exaggeration. I mean a literature search is might be something you ask a PhD student to do. It's more focused than that. You know I will pick out two or three papers that I think are the most important and of course they can lead you on to other... two or three other papers and you'll end up reading a dozen papers, but I wouldn't try to be comprehensive in my literature search. [AAA4]

I mean if you have got a particularly new subject or you try to fill in the background on some subject and it is a literature search or, I mean I find nowadays I don't tend to do much my literature searches. I usually hire the students who are

working with me doing those, or, or we have a series of papers anyway. We follow that references from one paper to see where the, what fell into that paper usually, so it is not often now that I actually get to go to the web and actually do literature studies. [AAP1]

### 5.2.2. Research Students

Physics and astronomy are expensive sciences and especially in experimental areas, research students join a team and work on a part of a project. The following quotation from a research student in astronomy shows how students integrate in their research group and environment and how they start and choose their PhD project.

When I first started, I did some small projects for the first 6 months. I helped them with some data analysis and data work; you do a couple of presentations. A sort of getting an idea of how people around you working because you're going to be working with them and generally I got interested in depth in certain areas and then I really didn't have a sort of research into any potential projects at that time. And then around about March my supervisor said it was probably a good idea to choose a project, back then he offered me 2 or 3 and I thought about them and did some reading of papers by a sort of people who've done previously in the field and then I chose the one I was doing and a sort of do a lot more research then obviously when I was starting the project. [SAA1]

Supervisors and colleagues play an important role in the integration of research students in the department's research environment. Different supervisors may have slightly different styles in their supervision. However, the basic general process of supervision is the same in physics and astronomy. Supervisors are the ones who introduce the students to the research and help them get started. They train the students both in terms of information-seeking skills and data analysis and problem solving. A professor described the way he would supervise his students as:

During that time [the time when students are doing the lectures] I would find them a problem that I was probably working on myself and I was pretty confident it could be done and I would work with them. So they would go away and learn up the techniques and the background and so on, but I would be looking very carefully after them. As I said, I would be pretty confident that it would be do-able within maybe a year or so. And then after a year or so they could write that up and that would form the first part of their thesis. For the second part I would give them a problem that I wasn't particularly working on myself, but I had an interest in. They would be much more (not completely) on their own. They would be expected to show more initiative and find the literature themselves, develop their own techniques and come out with ideas of their own. That would form maybe two third of their thesis. [AHE3]

During the early stages of their PhD, research students go through a training period with their supervisors and their senior colleagues, and they develop their information-seeking skills during the course of their studies. During the first year or first semester they attend some courses to acquire the background knowledge they need for their research. They also learn

how to work with data and tools and how to fit in a research environment by doing smaller things for other projects that are run by their research group. The supervisors, as it appeared from the interviews, play rather a proactive role in training students for seeking information. The supervisors tend to delegate some information-seeking tasks to their students. However, this is not to merely reduce their own work-load, but to train their students how to find information. A supervisor in high energy physics explained why and how she delegates information-seeking tasks to her students:

...because I want the student to know about a resource and so I'll tell him to go find something and then when they can't find it I'll say well did you look in this place and so I found that's worked pretty well. Other times if it's something... well, mainly it's because I want them to know about a resource I think that's when I'll delegate. [AHE1]

A student in astronomy described how he got introduced to information sources in his field by his supervisor and his colleagues.

From the very beginning I was recommended to use NASA ADS for browsing astronomical journals (by my supervisor and other postgrads), and this is the main database I used and still do! [SAA3]

The common way in which research students go through the process of initial familiarization with the research subject is to get introduced to a few key resources, usually papers and sometimes books or conference proceedings, by their supervisors as a starting point. In theoretical physics these are likely to be classical review papers on the subject while in collaborative experimental physics they tend to be collaborative papers produced by group members just for internal use, as well as journal papers. These are some quotations from students, showing how they started their research.

...my supervisor got me started with some background reading by providing a couple of review papers from astronomical journals, and I did cross-reference from these. [SAA3]

...the way they came to me was from my supervisor. He was the one who provided me with the information that I required, papers and etc., he gave me a list of things that I should read to get started. [SAM1]

At the start I didn't do that much, as much as I should have probably because, my supervisor gave me a bit of background information and a little code to play with so I did that and I tried to understand what it was about and a sort of read general background information on that topic. [SAM3]

The other important source of information for students at the beginning of their studies as well as during the course of their PhD is talking to senior colleagues in the department. Fellow research students, researchers and other colleagues could be very helpful. This is specially the case when there is physical proximity and people have easy personal access to each other in the department. The friendly relation in the department has been improved by the



informal meetings that some of the research groups hold for chatting research. For example the astronomy group holds a very informal meeting called Astro-ph Pint normally every Thursday evening during which people bring along an abstract or two from the Astro-ph e-print server for informal discussion over a drink. They also use this opportunity to discuss their research ideas or problems with one another afterwards.

Once the students have got acquainted with the research project and some important background information about their research project, the next step they would take is to look up the references of those papers for finding more relevant materials and get more familiar with the research area.

Doing a comprehensive literature search at the beginning of PhD by using bibliographic databases seems not to be the norm. This is partly due to the vague idea that students have about the focus of their research at the beginning. This makes it hard to do literature searches and therefore students start by chaining and tracking references.

It is worth mentioning that the doctoral programs in the United Kingdom (UK) are not run the same way as they are in some other countries such as the United States of America; therefore one must be cautious in generalising the findings to the students in other educational systems. The doctoral programs in the UK are different in that they usually do not involve coursework and students are immediately heavily involved in research. However, as stated by a professor [AHE3] in some subfields of physics such as high energy physics students who want to do a PhD after their undergraduate degree must do six to twelve months lectures because their undergraduate course is not sufficiently advanced enough to start a research degree. The following quotation by a research student in experimental particle physics confirms this.

Yeah, the way they do it here you have 3 months lectures. So you are still undergraduate student for the first three months and they give you 20 hours lectures a week and problem sheets and things like that. So that gives you theory and background and techniques to analyse large quantities of data and things like that. So you don't do any specific work maybe until January after you start. [SHE1]

The other point which is worth mentioning is that unlike research students in for example Arts and Humanities that usually start their PhD studies with their own research plans and proposals, as stated earlier research students in physics and astronomy (and probably some other scientific disciplines) tend to join a running research project and work on a particular part of it, or they are pointed to a research problem by their supervisors some time after the start of their PhD. It should be also said that although in theoretical physics it is possible that a student comes with his or her own research problem and starts working on it, it

is not very common. An example of this was the interviewee SCM1 who started his PhD by working on his own research problem in theoretical condensed matter physics.

### **5.3. Methods used for identifying articles**

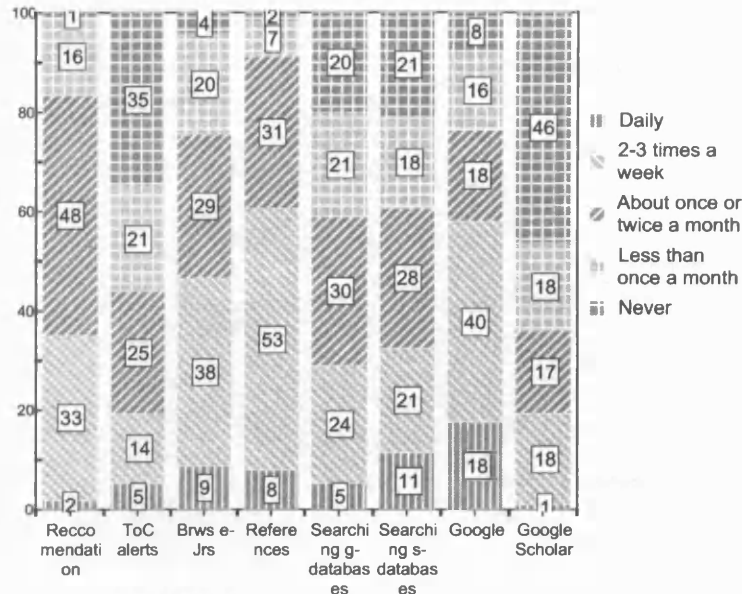
As Lynch (2007) said the primary vehicle for communicating and documenting results in most disciplines has been the scientific journal article, which has maintained a strikingly consistent and stable form and structure over a period of more than a hundred years now. Therefore, using journal articles is an important part of any research.

Scientists use different methods for identifying the articles they read. Sometimes it is a purposeful and active information-seeking activity such as conducting a search in a database; and sometimes it is a passive process such as receiving a recommendation from a friend or receiving email alerts. This section discusses the methods by which scientists identify articles they read. As this area is quite relevant to literature seeking, therefore, the author finds it appropriate to present this section here. The data is based on the questionnaire survey, in which the respondents were asked about the methods they used for identifying articles.

#### **5.3.1. By frequency of use**

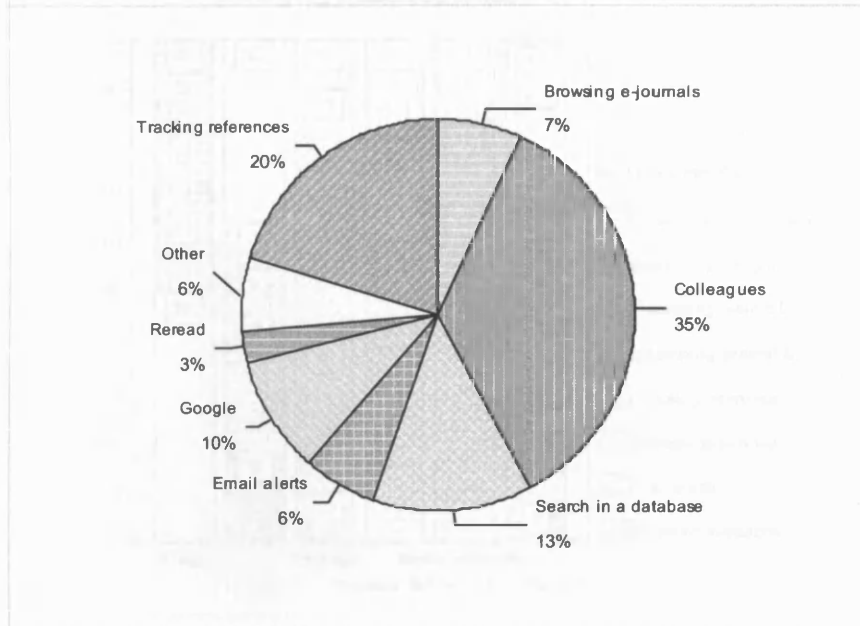
One of the questions in the questionnaire survey asked the respondents how frequently they used each of the presented methods for identifying articles. Regarding the frequency by which different methods were used for identifying research articles (Figure 5.1), Google stood on the top with 18% of respondents used it on a daily basis. Searching subject databases (11%), browsing or searching e-journal websites (9%) and tracking references at the end of articles (8%) were the other highly used methods on daily basis. Tracking references at the end of papers turned out to be the most popular method with 61% of respondents used it daily or 2-3 times a week, followed by Google (58%). 46% of the respondents never used Google Scholar for identifying research articles. This figure was 35% for ToC email alerts.

Figure 5.1: Percentage frequency distribution of methods used for identifying articles.



The respondents were also asked about the method they used to find the last article they had read (Figure 5.2). The most used method was recommendations by colleagues through which more than a third of the last articles read were found. Tracking references (20%) and searching databases (13%) were the second and the third most used methods. Although the previous figure showed that Google was the most frequently used method, this figure shows that a tenth of the last articles read were found through using Google web searching, which is still a considerable portion considering that Google is not designed for and is not meant to be used for finding scholarly articles. Only three percent of the last articles read were a reread.

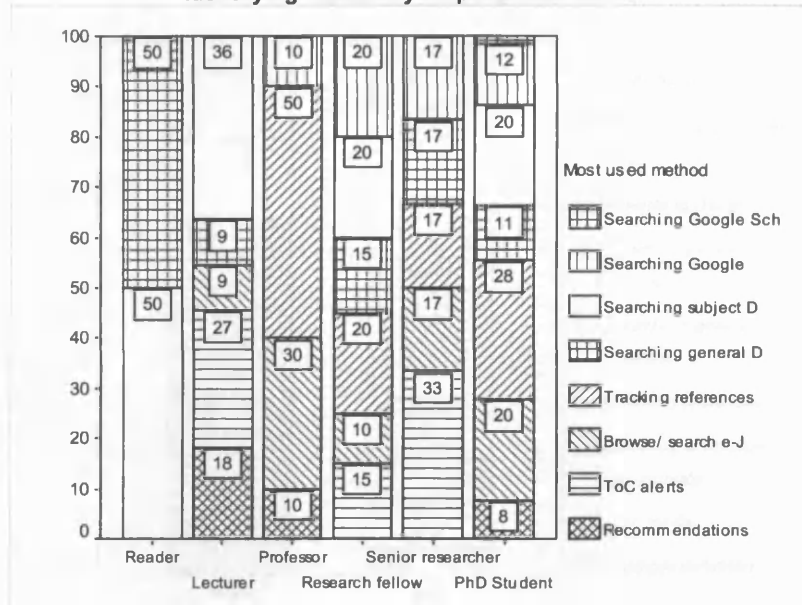
**Figure 5.2: Percentage frequency distribution of methods used for identifying last read article.**



### 5.3.2. By users' academic status

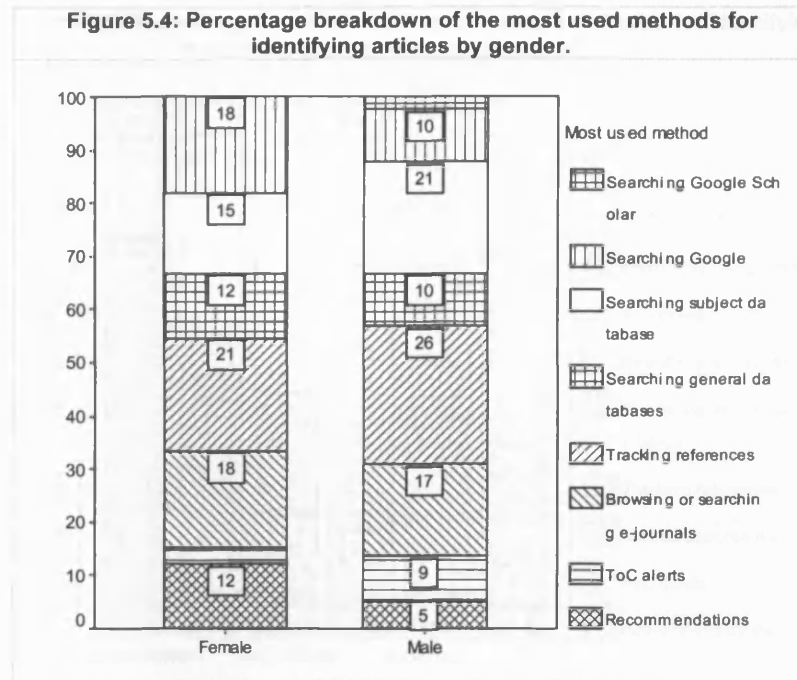
Another question in the questionnaire asked the respondents about the method they used most for identifying articles. There were some differences in the methods used for identifying articles among respondents with different status. Lecturers exploited 'searching subject databases' more than the other respondents did. 36% of lecturers used this method as their main method for identifying articles (Figure 5.3). Searching Google Scholar was not popular at all. Less than three percent of PhD students relied on Google Scholar but it has to be born in mind that the new studies (Hemminger et al., 2008) already show the increasing popularity of Google scholar since this survey was conducted. Tracking references at the end of articles was the favourite method for identifying articles for 50% of professors.

Figure 5.3: Percentage breakdown of the most used methods for identifying articles by respondents' status.



### 5.3.3. By users' gender

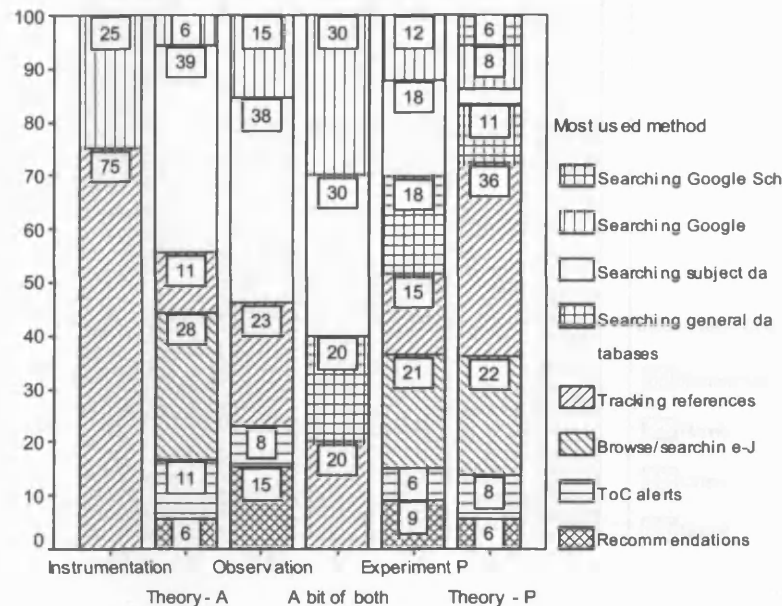
Slight differences could be seen between male and female respondents in terms of methods used for identifying articles. Female respondents were more reliant on recommendation than men were. None of the female respondents used Google Scholar to find articles as their favourite method. On the other hand, 21% of men chose 'searching subject databases' as their most used method for identifying articles compared to 15% of women (Figure 5.4).



#### 5.3.4. By users' type of research

Type of research refers to whether the research conducted by the participants was experimental, theoretical, observational and so on. This classification of the research in physics and astronomy was designed by consulting the literature as well as talking to a few professors in the department (they were among the interviewees) about how the research in physics and astronomy could be categorized. Looking at the type of research and the most used method for identifying articles we could see a high reliance of instrumentalist physicists on tracking references. 36 percent of theoretical physicists also relied on this method for identifying research articles (Figure 5.5). Searching subject databases was the most used method for theoretical astronomers, observational astronomers and those physicists whose research combined both theory and experiment. This difference can have different reasons including the availability of the databases in the field as well as other reasons. For example the high reliance of theoretical physicists on reference tracking might be because it is an efficient method for tracking an idea or a theory in the literature. However, this is an area that merits further investigation.

**Figure 5.5: Percentage breakdown of the most used methods for identifying articles by type of research of respondents.**



### 5.3.5. By interdisciplinarity of the field

Respondents were asked to specify how often they needed to use the literature of other disciplines, implying how interdisciplinary their research was. Of course this is based on their perception of their field and may not a hundred percent match the reality. However, the assumption is that they know their field and know how much they rely on the literature that belongs to other scientific fields. Figure 5.6 shows how often respondents in each group used literature of other fields. The data show that groups such as CMMP and TMP are more reliant on the literature of other fields and therefore have a more interdisciplinary nature while other groups such as HEP rely mainly on their own literature and cannot be counted as interdisciplinary.

The cross-tabulation of the interdisciplinarity with the most used method used for identifying articles (Figure 5.7) shows that those with the most interdisciplinary research i.e. those who used the literature of other fields often, relied on Google scholar (100%), ToC email alerts (63%) and searching general databases such as Web of Science (58%). On the other hand those who used the literature of other fields never or rarely tended to rely more on the recommendation (63%), searching subject databases (54%).

**Figure 5.6: Percentage breakdown of use of interdisciplinary literature by**

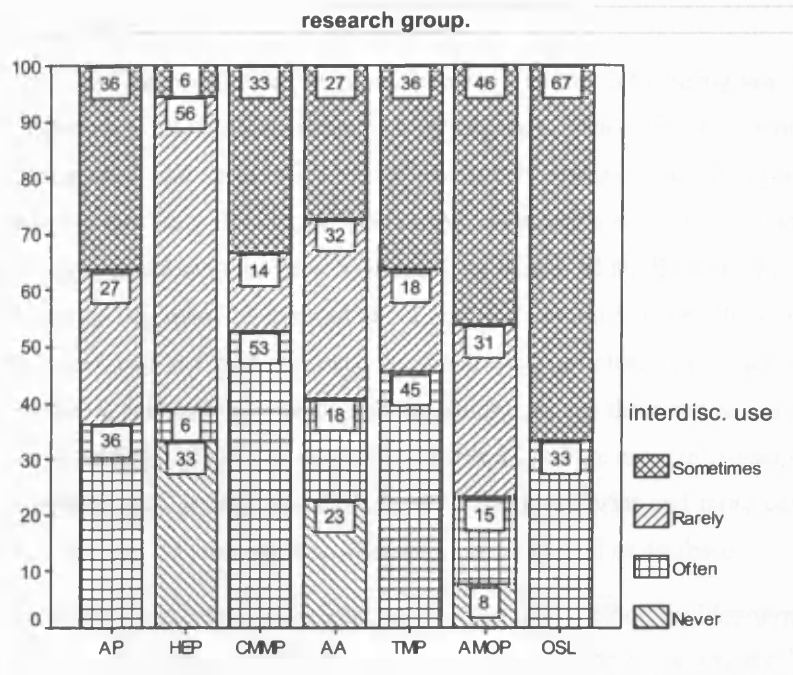
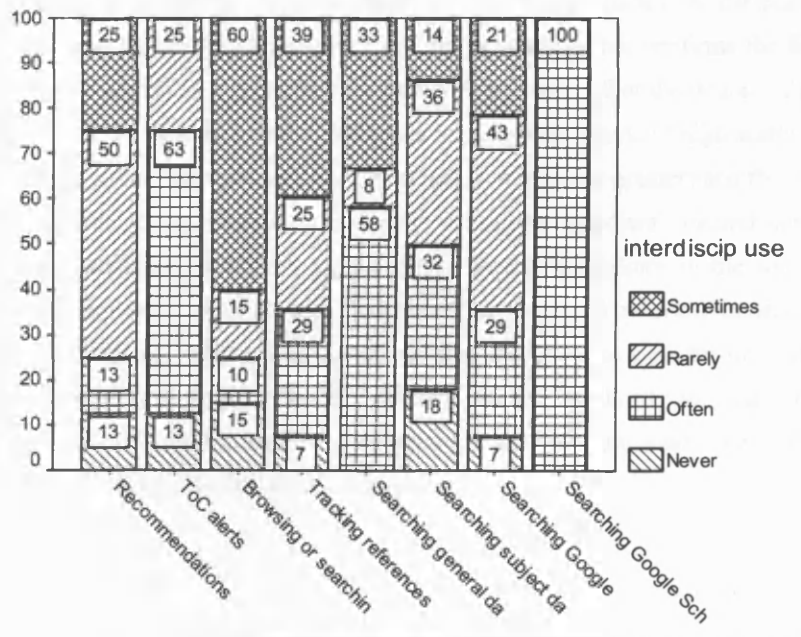


Figure 5.7: Percentage breakdown of the most used methods for identifying articles by interdisciplinarity of the field.





### 5.3.6. By scatter of the field's literature

Bates (2002:138) has suggested that one of the important factors influencing search patterns and the use of electronic information resources is the degree to which information on a subject is distributed (scattered) among the resources where such information may be expected to be found. In order to see how this factor affects the information-seeking behaviour of the participants, respondents were asked how scattered they perceived the literature of their fields to be. This question was asked to find out about any relationship between the scatter of the literature of a field and the methods used for identifying articles. The characteristics of literature in different subfields of physics and astronomy are not the same as we can see in Figure 5.8. Some subfields of physics such as CMMP and TMP are more interdisciplinary and some other subfields such as HEP and OSL have clearer boundaries and more concentrated literature in that they have a clear set of specialised journals as well as databases.

Figure 5.9 gives percentage breakdown of the most used method for identifying articles by the scatter of the subfields' literature. The more scattered the literature, the higher the likelihood of use of Google, Google Scholar, general databases and subject databases. On the other hand the less scattered the literature of a subfield was the higher the chance that respondents used Toc email alerts. This makes sense because databases cover several subject areas and they tend to be a better means for finding articles in several fields simultaneously. The cross-tabulation of the most used method for identifying articles by the scatter of the field's literature and by the interdisciplinarity of the field somewhat confirms the findings by Bates (2002) and Vakkari and Talja (2005). Bates (2002) argued that the degree of the scatter in a domain is likely to influence search strategies in systematic ways. In high scatter domains, the subject area is wider (the number of different research topics is greater) and the literature is less clearly organised or unhelpfully organised in the light of scholars' research interests and problems. Interdisciplinary fields are typically high scatter disciplines in the sense that the researcher must typically cross several disciplines to locate all relevant materials (Bates, 1996:156-7). Vakkari and Talja's (2005) study showed that the use of electronic resources and searching is more in fields with scattered literature. As the results of this study showed the areas that are more interdisciplinary or have more scattered literature are more likely to use general search facilities for finding their information.

Figure 5.8: Percentage breakdown of the most used methods for identifying articles by scatter of subfields' literature.

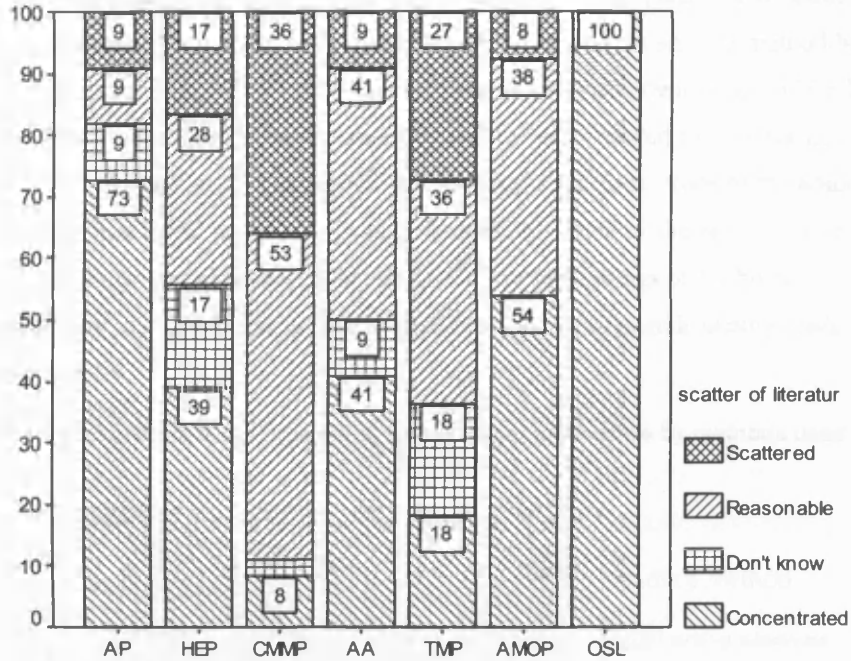
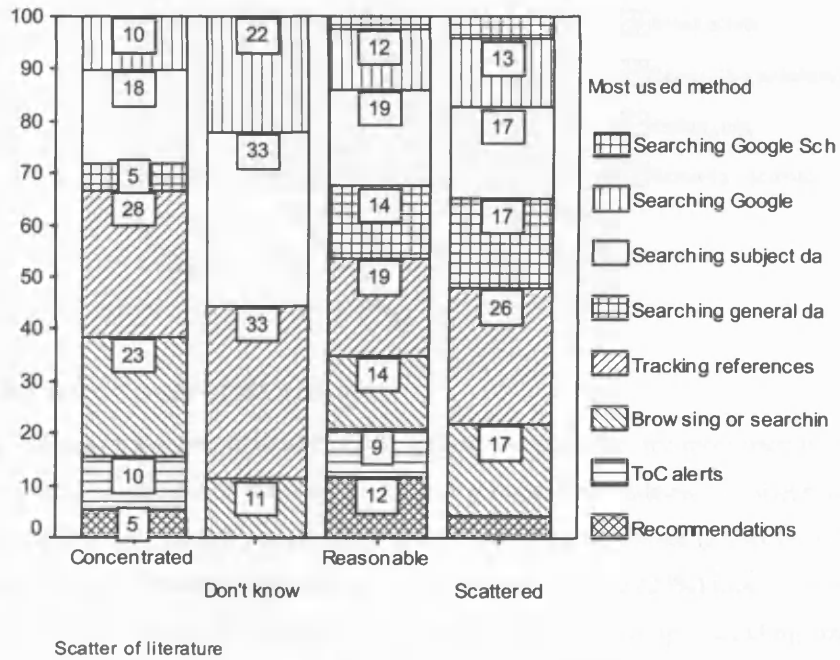


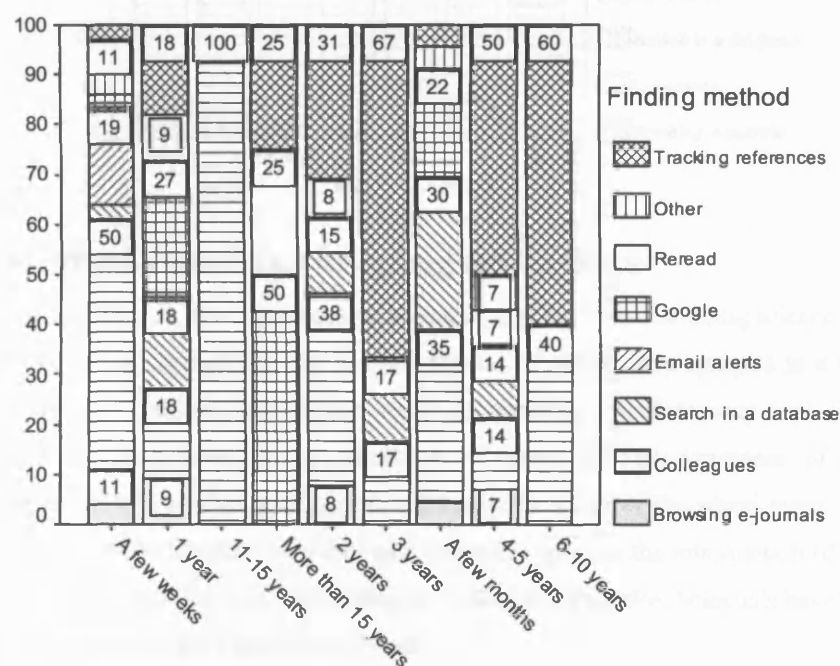
Figure 5.9: Percentage breakdown of the most used methods for identifying articles by scatter of subfields' literature.



### 5.3.7. By age of article

The questionnaire had a critical incident section that asked the participants about the last article they had read. They were asked about the age of the article and the method by which they had identified it. Figure 5.10 visualises the percentage breakdown of age of the last read article by methods used for finding it. Generally older articles tended to be identified through means such as colleagues' recommendation and tracking references. None of the articles older than five years was found through browsing. Surprisingly 50% of the articles more than 15 years old were found using Google. This results confirm the findings of Nicholas et al (2005a) and Huntington et al. (2006) that online availability and use of search facility leads to more use of older articles.

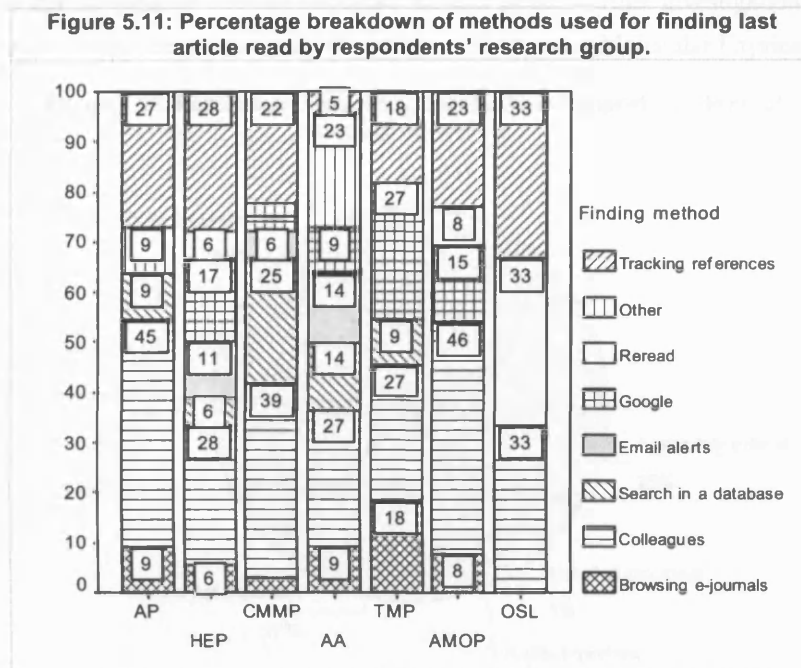
Figure 5.10: Percentage breakdown of age of last read article by methods used for finding it.



### 5.3.8. By users' research group

Tracking references and recommendations by colleagues were the two most used methods for identifying the last article read for most of research groups. Respondents in Condensed Matter and Material Physics (CMMP) made more use of searching databases (25%) than the other research groups did. Theoretical Molecular Physicists used Google (27%) more than any other research group did. Email alert was used only by three research groups including Astronomy and Astrophysics (14%), High Energy Physics (11%), and CMMP (6%). High usage of email

alerts by these groups can be attributed to the good email alerting services that are available in these areas through arXiv and the other sources. Figure 5.11 relates.



#### 5.4. Methods used for keeping up-to-date

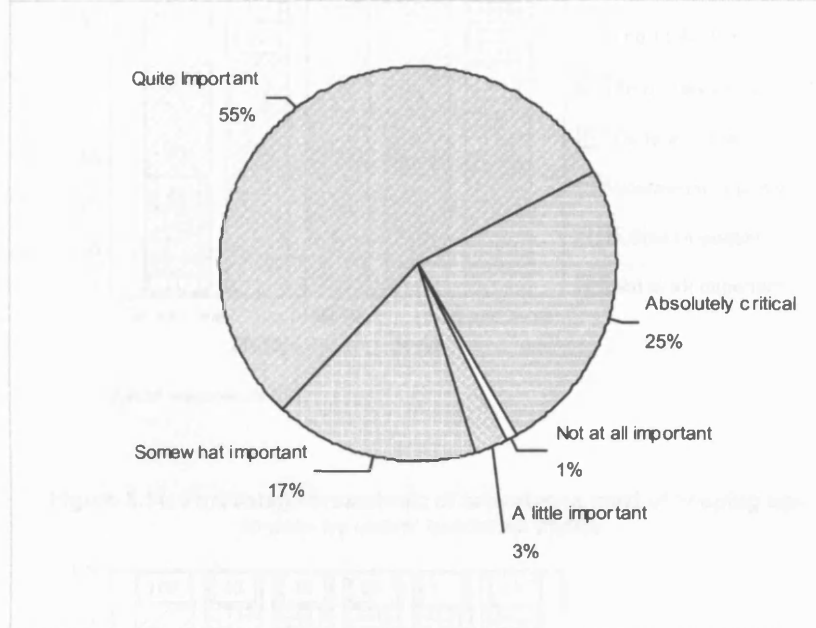
Keeping up-to-date with developments in a scientific area is crucial for doing science. Without having methods and mechanisms for keeping abreast of recent developments in a scientific area and without efficient information retrieval systems, scientists may end up doing redundant work or reinventing the wheel. As a matter of fact, ignorance of the past achievements and discoveries in science has led to reinvention of the wheel every now and then. Lancaster (2003) has discussed this with a few examples in the introduction of the third edition of his book *Indexing and Abstracting: in Theory and Practice*. Scientists have different methods and techniques for keeping up-to-date.

The ways by which physicists and astronomers keep up-to-date was one of the main issues that were investigated in this study as part of their information seeking activities. All of the interviewees were asked how they kept up-to-date with the developments in their field. Three of the questions in the questionnaire survey were also related to keeping up-to-date. This section discusses the findings of the study about methods used for keeping up-to-date.

As it might be expected from physicists and astronomers, the majority of respondents believed that it was important for them to keep up with the developments of their subfields (Figure 5.12). But the levels of importance were different. A quarter of respondents

considered keeping up-to-date as absolutely critical for their research. Fifty-five percent ticked the option 'quite important'. Although a very tiny minority, surprisingly one respondent maintained that keeping up was not important for him at all! Further investigation of the data showed that the respondent was a research fellow in Theoretical Molecular Physics.

Figure 5.12: Percentage frequency distribution of importance level of keeping up-to-date.



No specific correlation was found between age and importance of keeping up-to-date. Figure 5.13 does not show any meaningful relation in this regard. This is, as explained in the previous chapter, is partly because the sample is not an ideally representative one. However, we could see that the ones in category 34 and less were more likely to associate less importance with keeping up-to-date. Looking at the status of the respondents (Figure 5.14), it turned out that those who associated less importance with keeping up-to-date were more likely to be PhD students or research fellows. About three percent of PhD students and five percent of research fellows considered keeping up-to-date a little important and one research fellow maintained that keeping up-to-date was not important at all for his/her research subfield (Figure 5.3).

Figure 5.13: Percentage breakdown of importance level of keeping up-to-date by age of respondents.

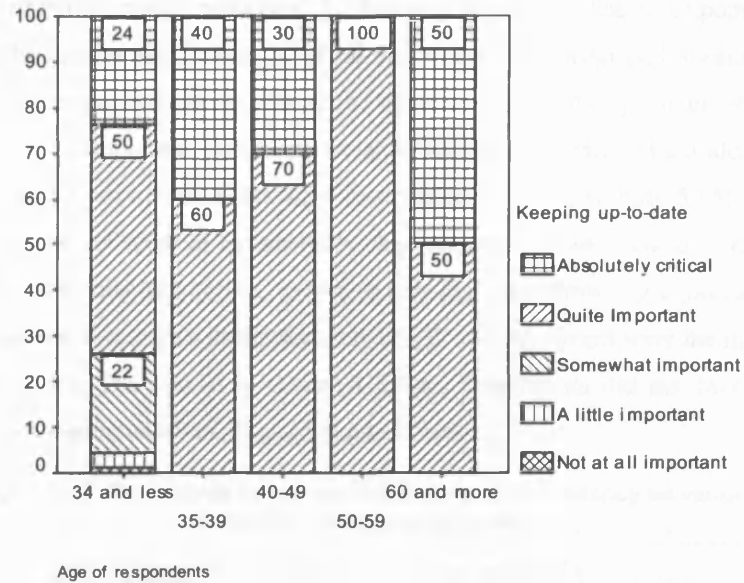
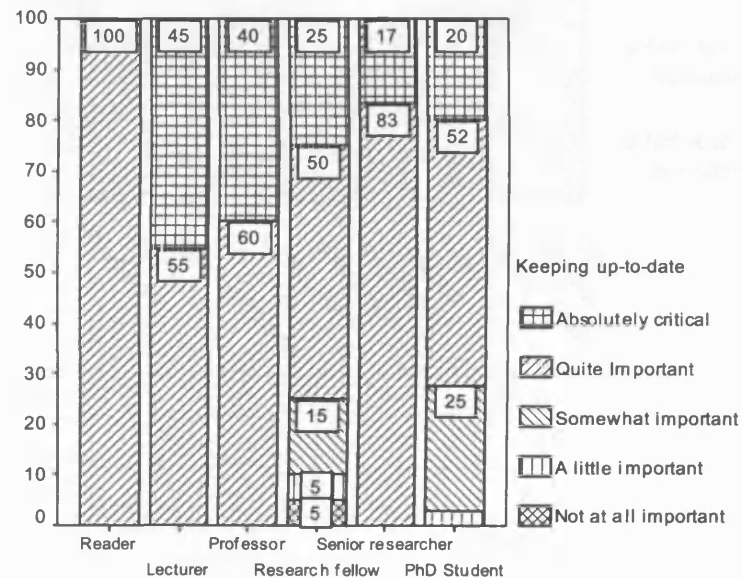


Figure 5.14: Percentage breakdown of importance level of keeping up-to-date by users' academic status.

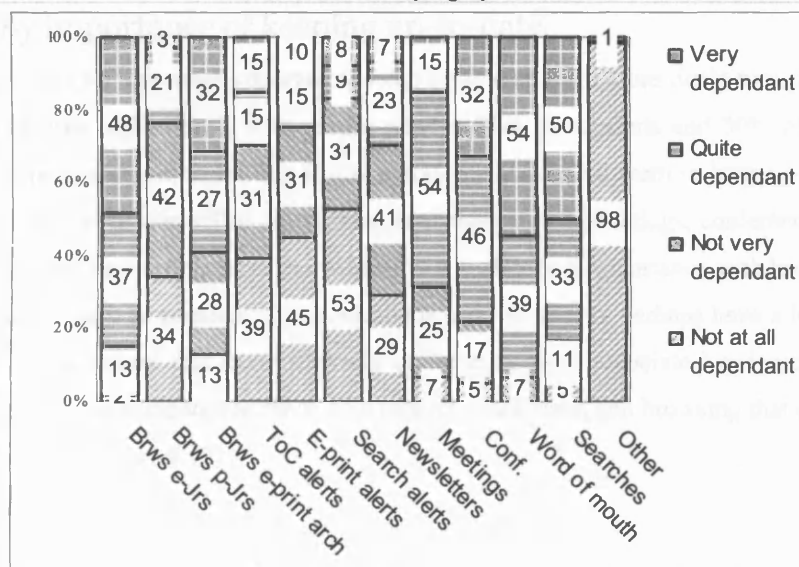


The following sections present the different aspects of the findings of the survey with regard to the methods used for keeping up-to-date.

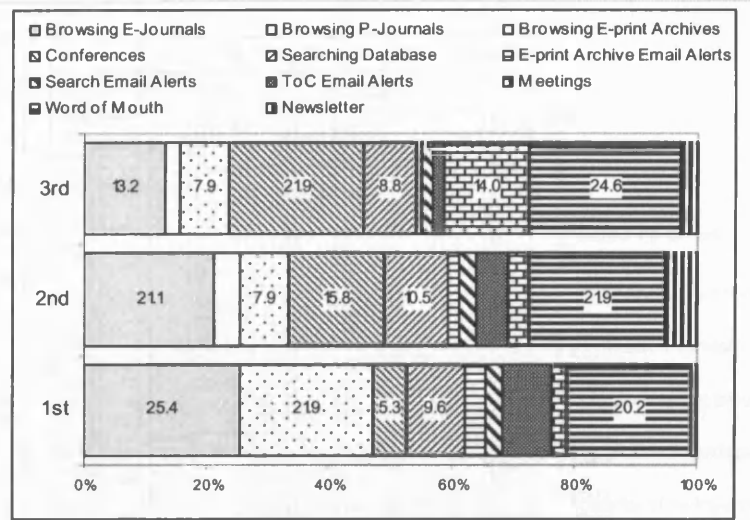
### 5.4.1. By dependency on various methods

A range of different methods were used for keeping up-to-date. The most popular methods turned out to be interpersonal methods. Word of mouth and colleagues, browsing e-journals, searching, conferences, and meetings were the methods on which respectively 93, 85, 83, 78 and 69 percent of respondents were very or quite dependent. Search email alerts, browsing print journals and e-print email alerts were less popular methods (Figure 5.15). This finding was also somewhat reflected in the answers when respondents were asked to rank their top three most used methods for keeping up-to-date (Figure 5.16). Browsing e-journals, browsing e-print archives and meetings with respectively 25, 22 and 20 percent were the most favourite methods among the first ranked methods. Although respondents did not favour receiving email alerts from e-print archives, they did browse them.

**Figure 5.15: Percentage frequency distribution of dependency on various methods for keeping up-to-date.**



**Figure 5.16: Percentage frequency distribution of the top three ranked methods for keeping up-to-date.**

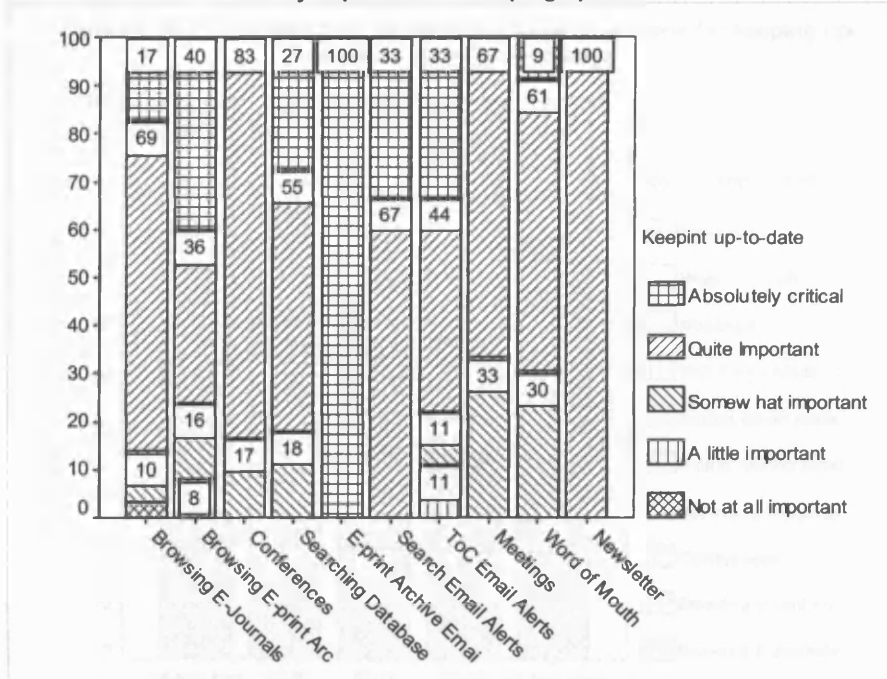


#### 5.4.2. By importance of keeping up-to-date

Those for whom it was more important to keep up-to-date were more likely to use e-print and alerting services. All of those who used e-print archive email alerts and 36% of those who browsed e-print archives stated that it was absolutely critical for them to keep up-to-date. On the other hand those who relied on personal communications (meetings, conferences, word of mouth) and also newsletters were more likely to associate less importance with keeping up-to-date. Methods such as personal communications and newsletters perhaps have a less frequent nature and it is natural that those who rely on these methods associate less importance with keeping up-to-date compared to those who rely on email alerts and browsing that can be done on a daily basis (Figure 5.17).



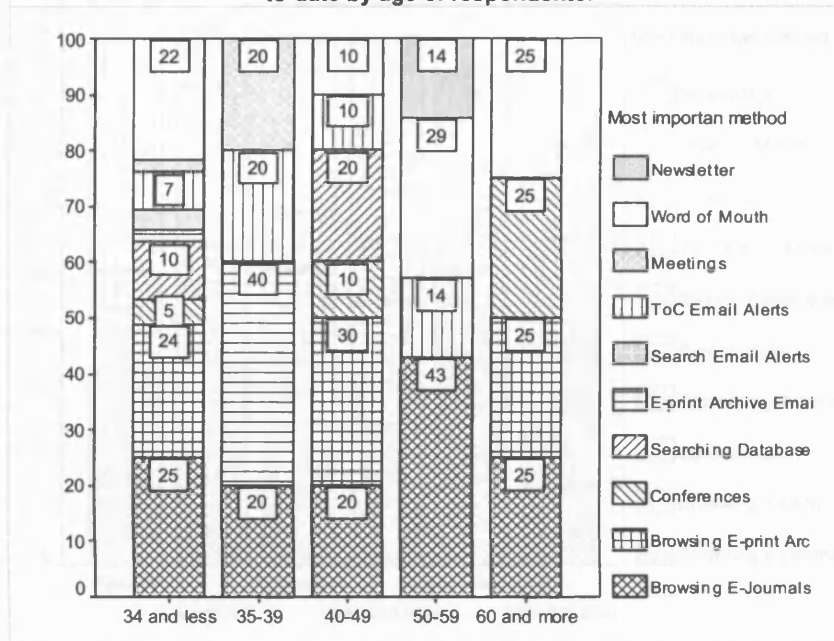
Figure 5.17: Percentage breakdown of the most used methods for keeping up-to-date by importance or keeping up-to-date.



### 5.4.3. By age of participants

People of different ages were dependent on slightly different methods. The oldest group (60 and more) depended on conferences more than the others did. A quarter of them ranked conferences as their first most used method for keeping up-to-date. Word of mouth was also an important method for keeping up-to-date for respondents above 50 years old. Perhaps it is expected that older academics rely more on personal communications such as word of mouth and conferences for keeping up-to-date compared to younger researchers. This is because older academics are expected to have a wider personal network and are more involved in their expert community through their longer academic careers while the younger ones may not have the privilege of access to a wide network of experts. Respondents between 35 and 39 depended more than any thing else on e-print archive email alerts, with two-fifth of them did so (Figure 5.18).

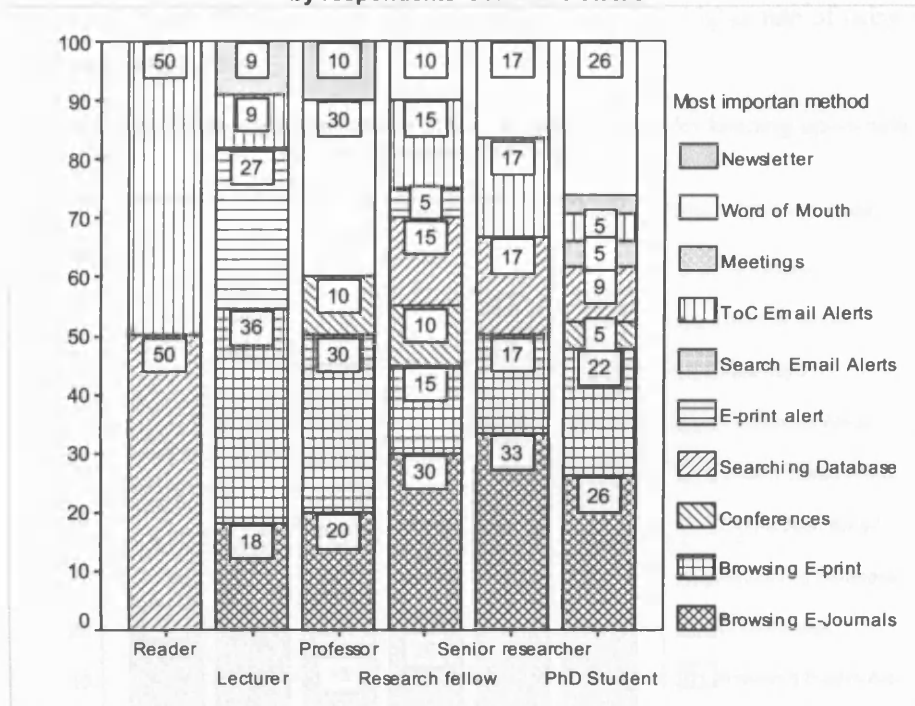
Figure 5.18: Percentage breakdown of top ranked methods for keeping up-to-date by age of respondents.



#### 5.4.4. By academic status of participants

Regarding the academic status of the participants (Figure 5.19), there were some interesting findings. Professors were the only group who chose newsletters as their top ranked method for keeping up-to-date, with a tenth of them doing so. Word of mouth was popular among professors (30%), PhD students (26%) and researchers (17%). Search email alert was not used and only 5% of students made use of it.

Figure 5.19: Percentage breakdown of top ranked methods for keeping up-to-date by respondents' academic status.

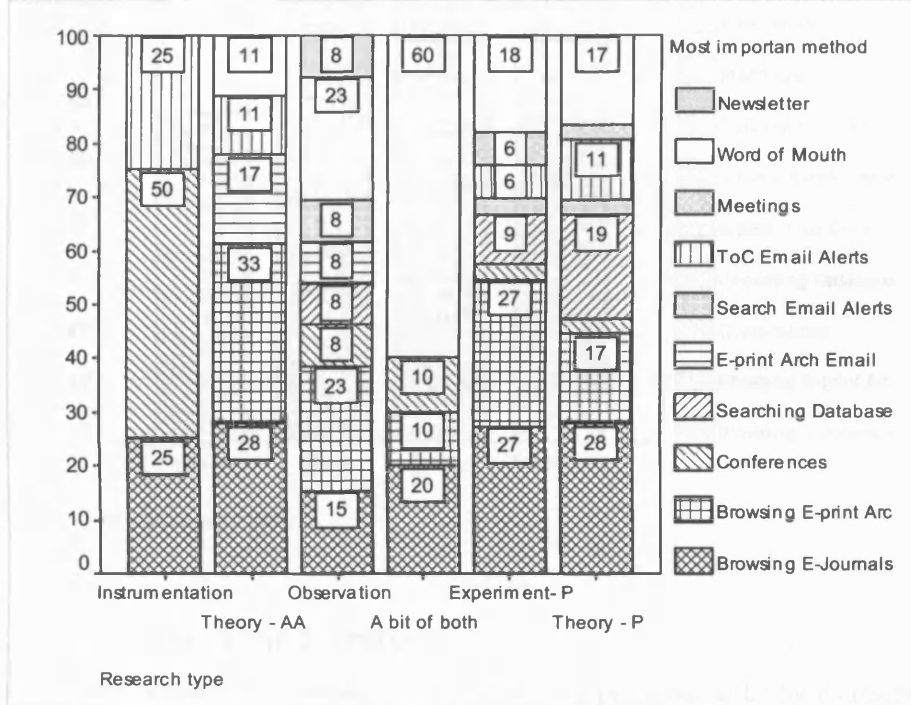


#### 5.4.5. By type of research

The breakdown of the top most used methods for keeping up-to-date with the respondents' research type (Figure 5.20) revealed the effect of type of research on the methods used for keeping up-to-date. Physicists and astronomers who were involved in instrumentation kind of research relied on conferences more than anything else, half of them chose conferences as their top used method. ToC email alerts and browsing e-journals (25% each) were the other two methods this group used as their most favourite method. Unlike the other groups who all browsed e-print archives as their first method for keeping up-to-date, this method had no use for instrumentalists as the first option. Theorist and experimentalist physicists resembled each other as they both relied considerably on browsing e-journals (28% & 27% respectively) and browsing e-print archives (17% & 27% respectively). However, they differed in that theorists made more use of word of mouth and meetings compared to experimentalists. Those whose research was a combination of both theory and experiment made the most of word of mouth, with 60% of them opting for it as their first priority for keeping up-to-date. Astronomers who were involved in observation also depended on word of mouth (23%), browsing e-print archives (23%) and browsing e-journals (15%). A notable difference between this group and the other group was that they were the only group who chose newsletters as their most

favourite method for keeping up-to-date, 8% of them did so. Theorist astronomers and astrophysicists' main difference with the other groups was their higher rate of using e-print archive email alerts (17%).

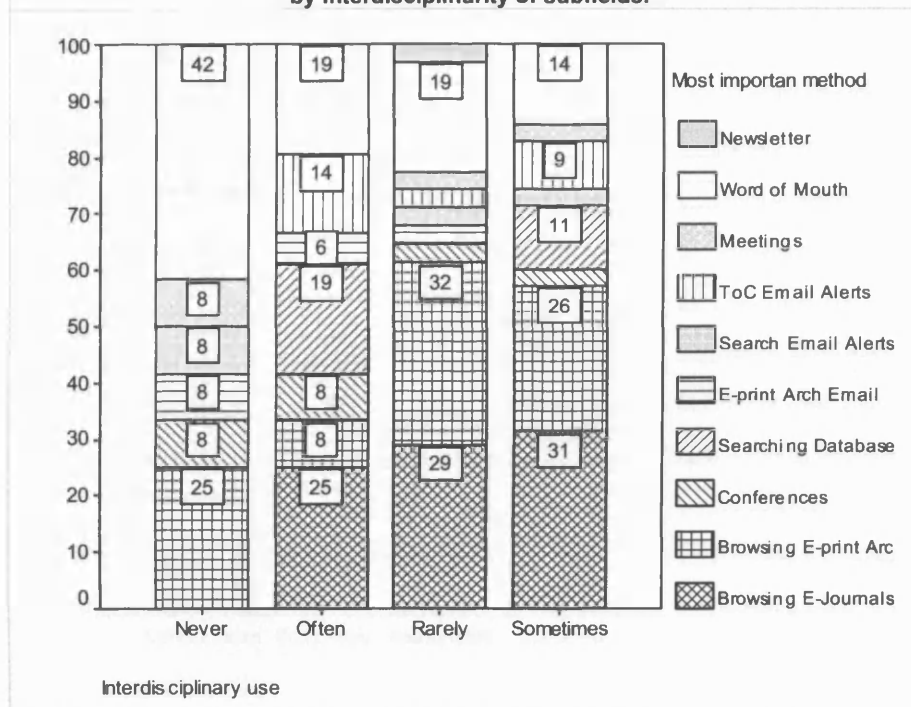
**Figure 5.20: Percentage breakdown of top ranked methods for keeping up-to-date by type of research of respondents.**



#### 5.4.6. By interdisciplinarity of the field

Respondents were asked to specify how often they needed to use the literature of other disciplines, implying how interdisciplinary their research was. Those with the most interdisciplinary research used 'database searching' more than the others. Nineteen percent of those who said they often need to use other disciplines' literature and 11% of those who said they did it sometimes searched databases for keeping up-to-date. The figure was zero for those who never or rarely needed other disciplines' literature. Figure 5.21 relates.

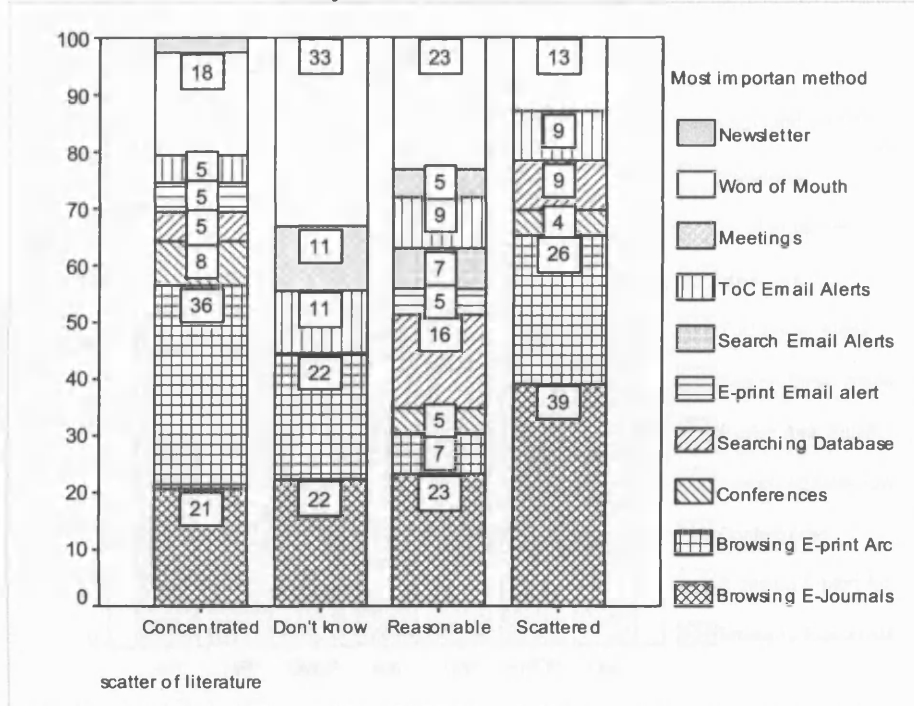
Figure 5.21: Percentage breakdown of top ranked methods for keeping up-to-date by interdisciplinarity of subfields.



#### 5.4.7. By scatter of the literature

Regarding how scattered the literature of a subfield was perceived to be by its researchers, those who believed their literature was more scattered were more likely to search databases and browse e-journals for keeping up-to-date. Those with concentrated literature had the highest rate of e-print archive usage (36%). Those who were not sure about the scatter of the literature of their subfields tended to rely on word of mouth more than the other groups did; a third did so (Figure 5.22). Again these findings are in line with the findings of Vakkari and Talja (2005).

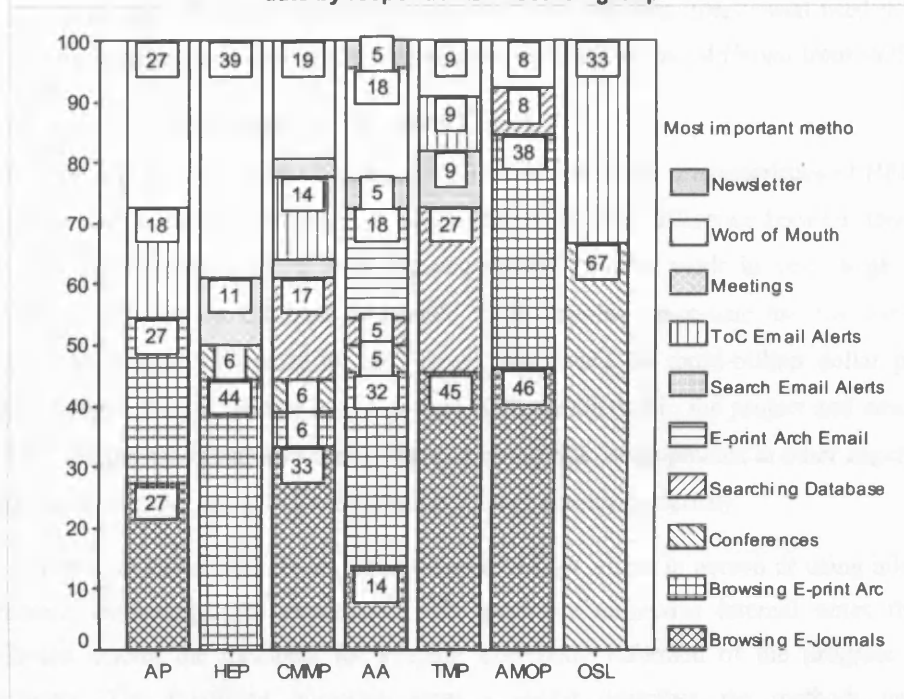
Figure 5.22: Percentage breakdown of the top ranked methods for keeping up-to-date by scatter of subfields' literature.



#### 5.4.8. By research group

Figure 5.23 gives the percentage frequency distribution of top ranked methods for keeping up-to-date by respondents' research group. Respondents from the Optical Science Lab relied a great deal on conferences, about two-third did so. Astronomers and Astrophysicists were the only group who relied partly on newsletters (5%) and e-print archive email alerts (18%) for keeping up-to-date. Five percent of them used newsletter as their main method for keeping abreast. The highest reliance on departmental meetings (11%) belonged to High Energy Physics. This group also relied on word of mouth more than the other groups did. Respondents in Theoretical Molecular Physics searched databases for keeping up-to-date more than any other group did. 27% of them depended on this method as their main method for keeping up-to-date.

Figure 5.23: Percentage breakdown of the top ranked methods for keeping up-to-date by respondents' research group.



#### 5.4.8.1. Atmospheric Physics (AP)

A major method for keeping up-to-date in Atmospheric Physics turned out to be checking the latest issue of journals either through browsing journals or receiving ToC email alerts. This was somewhat different from some other subfields of physics. The reason was twofold. Firstly, AP is not as large a research area as some other subfields of physics such as HEP and CMMP are. There are a few main journals that are active in the field of AP and they are known by people. Therefore they find checking these few journals relatively an easy and efficient way of keeping up-to-date. Secondly, while in some other subfields of physics, journals are considered out of date for the purpose of keeping up-to-date and reliance is on preprints, preprints are not very popular in AP. When asked about the use and popularity of preprints in AP, A professor in AP research group said that they were not popular in AP.

We have a lot smaller field, I mean we, the astronomers for example have a far larger number of papers published. They all sort of keep in touch with you to see who is publishing what. We have a very much smaller field. I mean probably four or five major publications producing in our field so it does not take a lot to keep up to date just by checking those as they come out and just checking the title, so we don't need to sort of elaborate procedures to keep up to date. [AAP1]

Conferences and word of mouth are the other methods that people in AP rely on for keeping up-to-date. The only research student from AP that was interviewed used the same methods for keeping up-to-date and his information-seeking was not different from staff's.

#### 5.4.8.2. High Energy Physics (HEP)

Reliance on e-print archives for keeping up-to-date is one of the characteristics of HEP. This is popular both in experimental and in theoretical HEP. The difference between theoretical HEP and experimental HEP is that experimentalists tend to work in very large groups sometimes consisting of hundreds of people. Their keeping up-to-date has two aspects of internal and external. Working in such large groups and on multi-billion dollar projects requires being aware of what is going on within the group and in the project and among the collaborators (internal) and also being aware of the recent developments in other experiments around the world and also of related theoretical developments (external).

Within an experiment, there are constant meetings, either in person or using telephone conference technology. An experiment also generates numerous internal notes that are distributed among the members for keeping everybody informed of the progress of the experiment. The following quotation from a reader describes the methods used by experimentalists for being aware of the progress of their experiments:

I mean within a given experiment there are constant meetings all the time, right, of the latest updates of results so within an experiment it's straightforward. It's just video conferencing to meetings, going to meetings or each experiment produces several hundred if not thousands of notes, internal notes, a year, which are on the web, which you read. [AHE4]

Newsletters and magazines are another important source of information for keeping up-to-date both internally and externally in experimental HEP. Important HEP facilities around the world such as CERN laboratory and FermiLab (Fermi National Accelerator Laboratory) publish regular electronic newsletters that contain recent developments in the experiments run using those facilities and also general developments in the experimental HEP. For example *CERN Courier* is an online magazine that is produced by CERN laboratory. Another example is *Long-Baseline Neutrino Oscillation Newsletter*. It includes rumours and references about neutrinos. The Newsletter itself insists that it should not be used as a reference, however, it publishes amusing anecdotes about different things that have happened in the neutrino world over the past month and then it also lists a number of references at the end that readers can check. This kind of rumour newsletter also reflects the importance of word of mouth which is another source of information for keeping up-to-date in HEP, especially in large research groups.



Journal Clubs and group seminars are also used in HEP for keeping up-to-date. A professor explained the use of journal clubs as follows:

We have a quite large experimental group here and they have a journal club. It's partly to explain recent things, but it's largely an educational thing for students to make short presentations. They would be told to go away, read up a particular recent paper and present it to the group in 15 minutes or something. So it's partly an educational exercise for them and partly for other people to keep up. There are also research seminars every week. Again they wouldn't be necessarily on what you are working, but they keep you up-to-date in the field as a whole. So all students would have to attend most of these. [AHE3]

The behaviours of research students in HEP in terms of their keeping up-to-date were not similar to each other. The theoretician student [SHE5] relied on the arXiv for keeping up-to-date. One of the experimentalist students, who was based in laboratory outside London and was email interviewed [SHE6], relied on meetings for keeping up-to-date. This might be because of the laboratory environment, something that needs further investigation. Two of the other students mentioned Spires database as their main source of information for keeping up-to-date. They checked Spires regularly. Another experimentalist student relied on one of his colleagues in his research group. This was an example of collaborative information-seeking in HEP group. When he was asked how he kept up-to-date he replied:

I'm quite lucky actually there is one person in our group who searches the databases everyday and send me the new developments and papers. The research I'm doing now there is a group of ten of us and one of these people looks through everyday and he sends us all the papers that have been released which saves me a lot of time. [SHE4]

### 5.4.8.3. Condensed Matter and Materials Physics (CMMP)

CMMP is more interdisciplinary compared to some others subfields of physics such as high energy physics, atmospheric physics and astronomy. This interdisciplinarity was also reflected in the popularity of *Nature*, which is a general science journal, among members of this group. Three of the interviewees said they browse *Nature* (in print format) regularly for keeping up-to-date. *Nature* was also used by students in this group. Two of the interviewed students mentioned that they check *Nature*, preferably in online format. The other characteristic of CMMP was that e-print archives appeared not to be as popular in this area as they are in some other areas such as HEP, therefore the reliance for keeping up-to-date is more on journals rather than e-print archives. In this respect, CMMP somewhat resembles AP group. A few of the respondents in CMMP group were subscribed to ToC email alerts. One used the British Library's Zetoc service, which is a table of content service of electronic journals for the British higher education. Although email alerts were used by some of the members of this

group, there were also complaints about it. They believed that they are not as efficient a method as they wished it to be. Also one stated that he was subscribed to email alert

... but you know, we receive so many emails today which is sometimes I read it, most of the time I just delete it as soon as it comes. [ACM9]

A few others preferred browsing journals rather than receiving email alerts. After journals, word of mouth as well as conferences was the second most popular means for keeping up-to-date among members of CMMP group. Those who have extensive experience, wide knowledge of the field and are well-connected to the experts of their fields rely more on word of mouth and their personal communications than on the literature. A professor close to his retirement, who mostly relied on conferences and word of mouth for keeping up-to-date, when he was asked how he kept up-to-date, answered:

By keeping my ear to the ground. Again it's... again I'm untypical, because I'm you know, I'm reaching retirement age. I know the areas of science where I wish to maintain a presence and these are fairly tight areas where I know most of the major workers and so I will keep my ears open for papers which are or work which I know is coming up in those fields. So I probably learn more from going to certain key conferences than I do by checking the literature. [ACM1]

One of the research students relied on search of Web of Science. She conducted regular searches on her topic every few months in order to find out about the recent papers.

I think the main thing for that is to do regular searches on the web of science on your topic. If you don't do it regularly you don't actually see the new material that coming up. So you may have done it in the beginning of your PhD and then 6 months later there could've been a paper that you haven't actually read and you're working on something that's already been published. That's one main thing I think. [SCM3]

#### **5.4.8.4. Astronomy and Astrophysics (AA)**

In astronomy and astrophysics, the e-print archive (arXiv) is very well-advanced and the majority of the people in this area rely on the preprint for keeping up-to-date. As a matter of personal preference, some would subscribe to email alerts and some would regularly check and browse the abstracts and titles on the Web. A researcher said that he uses the e-print archive for keeping up-to-date and he described the process as follows:

I will receive like one or two e-mails per day from Astro-ph with all the abstracts of the new publications. In fact it's quite dispersive because it's throughout astronomy, so there will be cosmology papers that you know papers that don't actually apply to me, but what I do every morning I will get this e-mail and I will start going through every one of them. So maybe I will dedicate that one hour every day. I will read the titles and then if the title interests me I will read the abstract. If the abstract interests me, which is really rare, you know because there are so many like it would be a very small percentage of them. If the abstract I think is relevant to whatever I then go and take the paper and read the paper as well....At

that point I would say I would estimate that on average there would be one paper every day that I would read with some more or less some attention. But I would probably read four or five abstracts. [AAA7]

However, differences exist among subfields of astronomy and astrophysics in terms of their use of and interaction with the e-print archives. Experts in some subfields such as Cosmology make much more use of the e-print archive compared to some other subfields such as Stellar astronomy. This is because the speed of development in some certain subfields is faster or slower than some other subfields. The faster the rate of development, the more likely the reliance on e-print archives. A professor explained this.

What I do, as a stellar astro-physicist, stellar astrophysics is a well developed science, I mean we know the basics of how stars work. So I'm kind of looking at details and that contrasts with... the strongest contrast is with planetary scientists who have the sort of massive leap forward every time there's a space mission and the cosmologists who in the last 10 or 20 years you know have just been making big steps all the time. So for cosmologists, because the field is evolving fundamentally in a very rapid way, you know the first thing cosmologists will do in the morning is come in and check Astro-ph to see if there's anything new. The rate at which things evolve in the areas that I'm working is much slower, so I don't need to do that. So I guess mostly by networking is how I find out what's going on. By e-mail...conferences and e-mails and just swapping pre-prints with people I know. There's a newsletter in the area that I'm working, an electronic newsletter that comes around monthly. [AAA4]

As we can see in the above quotation, word of mouth and newsletters have more applications in areas where the rate of progress is slower. Reading newsletters is probably the second most important means for keeping up-to-date in AA. There are many newsletters that are published on specialised research areas such as Hot Stars.

Like HEP, journal clubs are also popular in subfields of AA. There are several journal clubs in different subgroups of AA group in the Department. The AA group also has some other semi-formal and informal meetings such as Astrophysics Pint that people in the department use for chatting with each other and informing each other of the latest news and developments.

Students in AA group rely more or less on the same methods for keeping up-to-date. They receive email alerts from Astro-ph, which is the astronomy and astrophysics section of arXiv. The slight difference is that students are broader in their subject when they skim the literature for keeping up-to-date and also their reliance on the Department's or group's internal meetings such as Journal Clubs and weekly seminars are more than the staff's reliance.

#### **5.4.8.5. Theoretical Molecular Physics (TMP)**

People in TMP group also relied a great deal on e-print archives for keeping up-to-date. A researcher [ATM2] said he browses arXiv about twice a month to check for the new papers in

his field, while another researcher [ATM1] said that he is subscribed to arXiv email alert service. Besides preprints, journals also seemed to be important for keeping up-to-date in TMP, unlike some other subfields of physics that considered journals out of date for the purpose of keeping up-to-date. A research student in this group also checks journals on a weekly basis. Of the two researchers from the TMP group who were interviewed, one said he browsed journals and the other was subscribed to journals email alerts. He also used some specialised email lists on AGB stars and Hot Stars.

#### **5.4.8.6. Atomic, Molecular, Optical and Positron Physics (AMOP)**

In AMOP, which is conceived to be competitive by its research community, the e-print archive is the main source of information for keeping up-to-date. Although it was appreciated by a researcher [AAM1] that there is no guarantee of quality, which can be social misconception and confusion sometimes, they need to know what exists in the e-print archive in order to keep up-to-date. Here also seminars and word of mouth are important means for keeping up-to-date. However, this depends again on the size of the group. The larger the research group, the more reliance on the word of mouth for keeping up-to-date.

Well I've got lots of seminars here so that is a good source of information, and talking to other people. This is very much dependent on the size of your group. The larger it is, you just, you can just care less and talk to people and you will come to know how it is. [AAM1]

Preprint was also the most used resource for keeping up-to-date among students. One of the students [SAM6] actually had set the relevant section of arXiv as the homepage of his browser, so every time that he turns on his computer in the morning and opens his browser the first thing he sees is the list of recent papers in his area of interest.

#### **5.4.8.7. Optical Science Laboratory (OSL)**

Research in OSL is instrumentation research. Research in this area requires a lot of engineering and technical information and it has multiple aspects. For example the relevant material for a researcher who works on building a new lens for telescope can be spread out in astronomy, engineering and optical physics journals. This makes it hard to keep up-to-date. However, based on the kind of the research that one is involved in, a few methods could be used for keeping up-to-date. A researcher relied on receiving email alerts from e-print archives and conducting irregular subject searches. A PhD student, who felt frustrated with his ability to find relevant information, relied on the general information on the Web and search in Google. This was because the information he required was the kind of information that

commercial companies would hold. Attending conferences also appeared to be a good way for keeping up-to-date in this group.

#### 5.4.9. Role of conferences in information seeking

Conferences play a crucial role in scholarly communication as they bring people with a common interest together so that they can share and exchange ideas and information. There are many international conferences in physics and astronomy besides the national seminars and conferences and institutional seminars. For example the Canadian Astronomical Data Centre<sup>13</sup> listed more than 250 international conferences and meetings for astronomy to be held during 2007.

Different people may attend and participate in conferences with different motivations. Some may attend because they want to network with other experts in their fields. For some others, presenting their work and getting credit and publicity for their research might be the main reason for attending a conference. Some find it a good place for getting inspired and discovering new research ideas. The initial sparks of many collaborative projects are produced at conferences. People may also attend conferences because they believe it is an efficient way of keeping up-to-date with the developments in their fields. People might also attend a conference for a combination of these reasons.

It is notable that different conferences may have different strong points and applications, hence different motivations for attending different conferences. A research student explained that he would attend different conferences with different motivations because different conferences have different goals and target groups.

The interviews showed that there were some differences in the role which conferences play in scholarly communication as perceived by different interviewees with different status. Research students and younger researchers at early stages of their career look at conferences as a good place to get to know the experts in their fields and to talk to people and have their opinions on their own research. This incentive is stronger in the case of students as they look for more support for and ideas pertinent to their own research. Also for those students who have just started their research it is a good opportunity to see the breadth of the subject area they are working in. Eight out of twenty-nine interviewed students mentioned this. A student for example stated that:

Because I just started my PhD so it gives you a breadth of the subject without as much of the depth as you may want in some of the things that people talk about.  
[SAM4]

---

<sup>13</sup> <http://www1.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/meetings/meetings.html>

Another research student for example explained that the conference would be an opportunity for him to find people who are interested in his research and talk to them and hear what they think about his research.

Well hopefully if I produce some good results I will put a poster up then anybody whose work is in the same area would hopefully come talk to me. Then it's really a good way of getting to know everybody in the field. At this stage I'm not heavily arguing or collaborating with lots of people. It's just a good way of seeing what everybody's opinion is on this subject. Because we all go to the same talks and if there is a talk hopefully and someone has an opinion on it then they will speak out and argue or disagree or something and that gives me a good feeling of how the field works. [SAA1]

As half of the research fellows mentioned, conferences are a good place for researchers who want to improve their career by advertising their research. For people with higher status and at older ages such as professors and readers and so on, conferences are a good place for keeping up-to-date. Normally, researchers present the findings of their studies long before they are published in journal articles. This method of keeping up-to-date has some advantages. First of all, the researchers are present and one has the chance of talking to them if any ambiguity appears. Conferences are the 'human side of doing science' as a professor [AAA5] put it. The research is also presented in a rather simpler language and less technically than in the language of published works. This makes it more pleasant for the audience who want to know about the research in the areas which are related to their own research but not directly so. A reader's remarks on the role of conferences were:

Oh well to, to get up to date with the latest developments. To find out what professor X is doing these days. It's a short hand version of the published work so rather than plough through a difficult paper, one might as well hear it in the flesh and understand it better and maybe avoid having to read the difficult paper. [ACM4]

A professor also highlighted the importance of conferences for keeping up-to-date and added that it is a good way of doing so because one has the chance to validate the information in a face to face conversation.

They're very important indeed because you hear what is going on, you hear developments going on, you get to talk to the people and so that in itself you know you can tell a lot from talking to somebody whether they know what they're talking about or whether they're just bullshitting. [ACM1]

Five of the interviewees among staff also mentioned the conferences as a place when the new research projects start to form. They said that some of the research projects come out of a chat at a conference, or are inspired by a paper that is presented at a conference. The following quotation which indicates the role of conferences for forming scientific collaboration is from a professor.

But the more important activity is really forming collaborators, forming new collaborations, informing people so new stuff can go forward. [AAA5]

Besides all of these objectives for attending and participating in conferences, the majority of both staff and students who were interviewed (about 80% of the interviewees) considered networking an important incentive for attending a conference. Of course forming personal networks is not the ultimate goal. It is important because it has benefits such as sharing information and expertise, forming research collaboration and opening the gates for informal communications. In a review of research on invisible colleges<sup>14</sup>, Cronin concluded that such informal scholarly communication networks are the 'lifblood of scientific progress for both the physical and the social sciences' (1982:225). Moreover, Cronin (1982) notes the following advantages of the invisible college in contrast to the more formal channels of scholarly communication: currency of information, specialization of information, opportunity for feedback and input at formative stages of idea development, and potential for interdisciplinary transmission of ideas. Some of these advantages were, as stated above, motivation of the interviewees for attending a conference. One must bear in mind that conferences are *formal communication*, but as Hensman (1977) stated these formal frameworks (conference, seminar and so on) often opens channels to more informal relationships between scientists. Therefore conferences are a good place to initiate informal communication channels. The interviews, as one would expect, showed that the communication is not restricted to face-to-face conversations at the conference venue. Once a connection has been established, new technologies such as email and the internet are used for communication among the members of the network which leads to what Gresham (1994) called 'Cyberspace College'.

Both staff and students interviewed mentioned how they utilised email and the internet as well as other communication means to maintain the connections they had made at conferences and similar events in order to get benefits such as: track down sources, and access information (such as papers) when they are not available at their own institutions; get feedbacks on their research; share subject specialties; exchange and critique pre-publication papers; form collaborative research and writing; learn about job or research opportunities; calls for papers and so forth.

#### 5.4.10. Role of research environment and colleagues in information seeking

Research environment is an important factor that can influence the information seeking and communication behaviour of scholars. People are one of the building blocks of the research

---

<sup>14</sup> Invisible College is a term coined by Price (1961, cited in Cronin, 1982) to describe the informal communities of scientific specialists.

environment. An informed colleague is an invaluable source of information. In some areas such as engineering colleagues are used as one of the main sources of information. For instance a research by Hertzum and Pejtersen (2002) concluded that engineers get most of their information from colleagues and internal reports. They investigated how engineers' information-seeking practices intertwine looking for informing documents with looking for informed people. This characteristic of information-seeking behaviour of engineers has a long history and had been identified before Hertzum and Pejtersen's study (e.g. Shuchman, 1981) and it has been interpreted as the Zipf's Principle of Least Effort (Pinelli et al., 1993).

People are a critical source of information because they can explain things and their knowledge is enriched by their past practical experiences. The interviews in this study also showed the importance of colleagues in information-seeking behaviour of physicists and astronomers.

Research students have expectedly close relations with their supervisors. And those whose offices are physically close to their supervisors' offices might meet them informally even on a daily basis. There is a great deal of information exchange between them and they point one another to new relevant papers and information that they come across. Supervisors sometimes also act as an authority for validating information, as students discuss any pieces of information or papers that they have doubt about their accuracy or quality.

This interpersonal information exchange is not limited to students and supervisors. Students also use the other colleagues, especially their senior fellow students or research fellows in the same research group as a good and handy source of information. They would refer to them and ask their questions. This is more likely to happen if they have a question or need some information that has a technical aspect or needs experience. For example if a student needs to know how to use a tool or a piece of software or to conduct a calculation, then their first source of information is more likely to be one of their colleagues.

In the information-event card study, 20 out of 88 (23%) information-seeking incidents entailed referring to a colleague. Research students accounted for 14 out of 20 incidents in which a colleague or a friend was consulted as an information resource and means. Members of staff accounted for only eight information events in which a colleague was used as a source of information. This shows the importance of colleagues as a source of information and reliance of students on this source.

The most important factor that has some effects on this interpersonal information exchange and use of people as source of information is the research environment in which



people work. In the case of the Department of Physics and Astronomy at UCL, a couple of factors appeared to have positively affected the interpersonal information exchange.

The Department has a couple of mechanisms for facilitating formal and informal communications between its students and staff. There are monthly seminars in which someone has to give a presentation about a research project. Some of the research groups also have a regular meeting called Journal Club. Someone reads two or three important recent papers on the subject of interest to the research group and presents their summary in the Journal Club and together they dissect the papers and discuss them. Some other groups have more informal meetings. For example the astronomy group has an Astrophysics Pint every Thursday afternoon, in which staff and students chat about their research, studies and papers over a drink.

The fact that the Department has tried to remove physical barriers as much as possible has also helped improve the communication among people. For example all of the people in the astronomy group moved to a new building in 2005. Students' and staff's offices are all placed along sides of a corridor which also accommodate a common room that is used for informal meetings and chats by students and staff. A student [SAM4] in the Atomic, Molecular, Optical and Positron Physics group used the term 'bar based science' to describe the communication atmosphere in the department. He explained that people in the department talk about what they think and what they do in their informal meetings. The department has tried to improve inter-group and inter-department work and co-operation, so for example one group can actually apply techniques that have been developed in the other groups and bring them together. This is one of the reasons why UCL built the London Centre for Nanotechnology, where people from chemistry, math, biology, computer and physics can talk to each other in the same environment. The same student maintained that:

where there is potential for a synergy for cross over people tend to find each other in the most parts. But there are cases where people came over and said oh I'm doing this as well but slightly differently in a different field or people say oh that's interesting, I think it's quite organic, it just happens. But there are different systems becoming more common in the department which help that kind of things go on.  
[SAM4]

The following quotation from a student also highlights the importance of the physical proximity among people from different research groups, for example different subfields of astronomy and astrophysics. This makes it easy to get acquainted with research in peripheral areas and see how people in other relevant research areas may look at your research from a different perspective. It is also an effective way of keeping up-to-date. Past research has already established the importance of physical proximity in scientific communication. For example Kraut, Egido and Galegher (1990:2) showed that physical proximity helps scientists

avoid or minimize many of the problems that arise in the process of conducting research, meeting partners, defining problems, planning projects, supervising coworkers and subordinates; and it also leads to collaboration because it is likely to lead to informal communication.

It's got a lot better now that we are down here. Because we just moved to these new offices and all the astronomers are in the same place so we've got this room {referring to the common room} and we can actually talk to each other. It's very easy to get stuck in very specialised groups. And as my research crosses several it's interesting to see who talk to who, it tends to be very close often you know we do star formation so we talk to each other, rather than seeing people who do Galaxy and do the same thing at larger scale or whatever. But the fundamental thing to me is communicating about research. [SAA2]

Another student also was pleased by the ease of communication, not only within the Department of Physics and Astronomy, but also between the Department and other departments such as the Chemistry.

...when I talk to people in chemistry area of physics I found it very good. A couple of people I know there and we see their results and they produce data on chemical substances and we use that data in our models. I found it very good. Everybody is willing to share the information you know. They trust their data and they want people to use and the more people use it the better they look. [SAA1]

Use of people as information sources is not restricted to people within the Department and people in the vicinity. Modern communication especially email has shortened long distances. Students mentioned for example that if they could not obtain a paper online and through UCL library they might try and contact the author for a copy of his or her paper. If they read a paper and need further data from the research presented in the paper they would ask the author. They would liaise with the experts in their research area through conferences and seminars in order to get advice for their research.

In the case of staff, they also rely on each other's knowledge a great deal. As explained before, both for staff and students, colleagues and generally people are a good source of information when they want to start a new research. A research fellow stated that:

Yeah, I, well so you know one way is, is that it... when you are working in a large group like, like we do, there are you know people who have known expertise on a certain particular little topic, right, so if you want to know about that little topic you would just e mail them or call on the telephone. Yeah and you know, they may know the answer or they may know that, you know, some, some work that you should go read to learn more about, about a particular topic so, you know, one takes advantage of ones colleagues in this way, right and so people will take advantage of you as well, right. I mean that's the whole point of us all being here together in this department, right. [AHE5]

A professor [AHE3] maintained that the role that colleague plays in one's information seeking activities somewhat depends on the size of the research group and the atmosphere in

which one is working. While in larger groups people make more use of one another, in smaller groups use of web-based facilities and libraries is much more important. This is the situation in particle physics. While people in the experimental particle physics group that is a large group talk to each other a lot, people in the theoretical particle physics group, which is small, rely more on the use of web-based resources and services. The professor maintained that going to a colleague for information would be the first choice if the right person is available and the atmosphere is suitable for that type of communication.

Use of technology for personal communication and using people as sources of knowledge has expanded due to the increased popularity of email discussion groups and collaborative authoring tools such as Wiki technologies. A senior researcher mentioned use of Wikis and instant messengers by him and his colleagues in the CMMP group both for collaboration as well as sharing information.

...we often will have meetings for some reason or other and we will discuss a subject and people's knowledge will be shared just around the table. That is very productive. Sometimes it is shared ignorance in which case it's not so productive. People outside the room then electronic media consists of discussions, either through Wiki or e-mail, instant messenger, telephone and those discussions will either be... a lot of it is just going through a problem. There is no reference to other papers necessarily or the work. It is just trying to work through the logical problem. That is a large part of it. Then we'll just simply send e-mails out to each other if you find something very interesting and sometimes we'll post the PDFs on the Wiki and say take a look at this. [ACM10]

The important outcome of the study with regard to the role the research environment and colleagues play in the communication and information-seeking behaviour of physicists and astronomers is interpersonal communications. Historically oral communication has been a heavily used but seldom documented method for communication in informal physics (National Research Council, 1973, cited in R. S. Allen, 1991:28). The most basic form of oral communication is informal personal interaction. According to R. S. Allen (1991) it is quite natural that scientists who spend much of their time conducting research will tend to speak with colleagues about that research. This type of communication takes place wherever the scientist is, be it in an office, hallway, restaurant, library, gym, etc. As most of the formal restraints common to many other communication modes are lacking, this is an opportunity for physicists to let their creativity flow, (Kasperson, 1978) and to trade bits of scientific gossip (Traweek, 1988).

Although new communication technologies such as email and instant messaging have affected the scholarly communication, the study's results show that oral communication still is a very important element in the scientific communication, shaping collaborations, initiating research projects and most of all seeking information.

## 5.5. Problem-specific information seeking

Information seeking is not only about keeping up-to-date or systematic literature searching that a researcher may do at the beginning of a project. A user may face an information problem, or in other words have some sort of uncertainty caused by a gap in his or her information that can be resolved by obtaining the appropriate information. This kind of problem-specific information-seeking<sup>15</sup> activity accounts for a considerable portion of one's information-seeking behaviour. This has been investigated through interviews and also in the information-event card study. When the interviewees were asked at what phase of a project they search for the literature, although the main answer was at the beginning and close to the end of the project, seven of the interviewees also mentioned that sometimes in the middle they might need to search as well. These are the cases when a researcher encounters a problem and needs information for solving it. The following quotation provides an example of a search for literature in the middle of a research:

Let's say I work in astrochemistry. So I look at chemical evolution of star forming regions. So in some cases there could be a problem that for example I do the simulation of this chemical model and I find that a particular molecule for example is way too abundant with respect to what is observed. So if I want to search recent XX in particular region of space where the molecule is abundant then I may try to see whether in the literature other people have found the same problem. So that would be more or less in the middle of doing the project. [AAA1]

Those who conduct experimental research or are involved in research with technical nature are more likely to seek literature in the middle of a project. Experimentalists need to seek literature in several stages. One stage is the interpretation stage. Once an experiment has produced some results, researchers need to look at theoretical literature to interpret their findings.

Well because we generally do fairly focused experiments, when we get to the interpretation stage with the data, we would then be doing a lot of searching primarily to see whether what we've got is actually as new as we think it is. In some cases also to get further theoretical background sometimes because as we get more sophisticated in the experimentation, the sort of level of interpretation requires quite a lot of theoretical background. Some of that we get from the models but again you, you know, you need to understand what... why the model is doing a certain thing so some of it is to expand that and get contextual information and other times it is specifically to compare to other experiments but certainly at the interpretation stage. [AAP2]

Experimentalists or those with technical kind of research such as instrumentalists may need to look for technical information in the literature during the course of a research project. For example if they take some measurements they need to look at literature to see if anyone

---

<sup>15</sup> The term problem-specific information seeking has been borrowed from Savolainen (2008).

else has done another measurement and compare their measurements with the others'. An experimental high energy physicists [AHE4] raised this point and stated that once a measurement has been done and is to be published, comparisons must be made with previous measurements or the measurements by the competitors and these need to be referenced.

In the information-event card study participants were given booklets to record their information seeking events in a week time. The information seeking events were not restricted to specific types and they could include any kind of actively seeking information from any source. In total 88 information seeking events were recorded by 27 participants. As the following table (Table 5.2) shows more than half (56%) of the information seeking events were fully successful as participants found all the information they looked for, and another 20.5% were partially successful as the participants found some of their needed information.

**Table 5.2: Distribution of the information seeking events by their outcome.**

The outcome of information seeking event	No	%
Yes, found the information needed	50	56.8
Yes, found some of the information needed and will continue searching	18	20.5
Yes, found some of the information needed and will not continue searching	10	11.4
No, no information was found	6	6.8
Other (for example one browsed just for keeping up-to-date)	4	4.5

The information seeking events could be categorised into two broad categories based on the type of the information sought (see Table 5.3).

- Unspecified information on a specific subject: this is when participants looked for general information (in the format of papers or any other format) on a particular topic for example looking for scholarly papers on a theory to know background information in order to prepare for a presentation. Sixty-four percent of all events (56 events) were of this type, of which 22 were ended fully successful as uses found all of their needed information, 18 found some information and said they would continue looking for information later on, six cases did not find any relevant information at all, and four cases found some information and decided to end their search and do not continue. Four other participants chose the option 'other' to explain the outcome of their information-seeking event.
- Specific information items: this is when participants knew exactly what piece of information they were looking for. These tended to be very specific and small pieces of information such as bibliographic information of a citation to make sure the reference at the end of a paper is written correctly, or the definition of word 'mixture' to use in a paper properly. Thirty-two percent of the events were of this type. Twenty-eight out of 32 cases (87.5%) of information seeking events in this

category ended with complete success as participants stated that they had found all information they needed. In the other four cases users found some of the information they needed and stated that they would not continue looking for the information. For example a participant who looked for a specific piece of documentary movie which he had seen previously could not find the same movie on the Web but found an alternative clip that he could use for his presentation.

**Table 5.3: Distribution of information seeking events by type of information sought.**

Type of information	Examples	No	%
Unspecified information on a specific subject	PDF papers about electrospin, background information for a talk, a set of recent papers on fluctuation theorem, degenerate electron conductivity data and any papers relating to this data, who is working in wave packet treatment of the TDSE,	56	64
Specific information item	Email address of a colleague, exact bibliographic information of a citation, melting and boiling points of potassium, definition of word 'admixture' and its difference with 'mixture', compressibility figure for hydrogen gas, a specific paper by a particular author that is about a specific argument, specific detail of a project by another research group	32	36

Table 5.4 gives the list of resources used for information seeking and the number of information seeking events in each case together with some examples. The total is not based on a hundred percent because more than one resource might have been used each information-seeking event. As we can see Google for general searching on the Web has been used the most, followed by using e-journals and asking colleagues or friends.

It is worth knowing that all who used printed journals and printed books, used books and journals from their own collections and did not go to the library for that purpose. All of those who used printed journals and the majority of those who used printed books (all except one who was a researcher) were research students. Research students accounted for 14 out of 20 incidents in which a colleague or a friend was consulted as an information resource and means. Among other highly used resources such as Google for general search on the Web, and e-journals, there was no significant difference between research students and staff, and they were, more or less, used by both groups.

Table 5.4: Distribution of resources and methods used in information seeking events

Resources used	Examples of information sought	No	%
Google	Email address of a colleague, fine information about molecular vibrational spectrum, search for Cost724, allowed electronic states for CrH, meaning and definition of word 'admixture', information about a new instrument	26	29.4
Google Scholar	papers or documents about buoyancy corrections for thermogravimetry	2	2.3
Web of Science	citations of a specific paper, publication of a recent paper by myself and colleagues	8	9.1
Checking e-journals	degenerate electron conductivity data and any papers relating to the collection of this data, looking for a paper I know the author and the argument from the references in another paper,	24	27.3
Checking print journals	information on some code for changing existing equation (already knew the relevant paper), information on computer code	4	4.5
Checking print books	expression for density distribution, background information for my talk, quantum chemistry information about numerical solution for Harper-fock	8	9.1
Checking e-books	Data on energies and electron status of HrC	2	2.3
Asking a colleague or friend	state of art of dissociative attachment calculations, calculation on tantalum, who is working in wave packet treatment of the TDSE, pump probe spectroscopy theory, find background information for talk, compressibility figure for hydrogen gas	20	22.7
Searching Hep-ex	general search of latest pre-prints in high energy physics	2	2.3
ADS abstract service	a paper on a particular observation	3	3.4
Wikipedia	melting and boiling points of potassium	2	2.3
arXiv	search for articles appeared in the month of January 2006	2	2.3
Searching Spire	exact citation of a reference	1	1.1

### 5.5.1. Web searching

What is evident from the information seeking events is the high reliance of the participants on the Web-based resources, and particularly general Web searching for finding information and resolving their information problems. As we could see about thirty percent of all the events involved using Google for finding information. It was mentioned earlier that Google was also the most used tool by which the articles that participants read were identified. This high reliance on Google and the role it plays in scholarly information seeking merits further investigation. Past studies have also raised questions on the use of Google by scientists. The results of the survey of physicists by CIBER showed the importance of Google. In CIBER's survey Google, after 'visiting a journals web site', was the second most frequently used method for identifying research articles (Nicholas et al., 2005b). This posed a question about Google. The question could be asked here is that whether the scientists intentionally use

Google for finding journal articles or they use it to look for any kind of information and as a result they are presented with journal articles among the results.

To cast some light on the role of Google, as the most popular general Web searching tool, the researcher extracted the comments that the interviewees made with reference to Google. The interviews of staff and students revealed that the participants in the study did not use Google to look for journal articles. The interviewees were specifically asked if they used Google for finding articles and the majority (except two interviewees) answered no. When they knew that they were looking for journal articles, they used scientific databases such as Inspec or ADS to find their needed literature. But a lot of time they do not have a clear idea as to in which format or source they might find the information they are looking for. It should be born in mind that the participants did not choose Google intentionally as a search tool for identifying articles, but finding articles is just a by-product of Google searching. They do not use Google as a tool for conducting a scientific literature search. The following quotation explains this. When one of the interviewees was asked whether he used Google for finding articles, the answer was:

No. Although I do find that increasingly when I Google for other things it does increasingly lead me to papers, which... the first time it happened I was very surprised. [AAA4]

This is the main reason why the results of CIBER's survey and the survey in this study show that a high percentage of the articles used by scientists are identified through using Google.

The preference for and use of Google has several reasons as it appeared in the interviews. The main reasons mentioned by the interviewees are listed here:

- Google is good as a start point and for getting a quick overview. Several interviewees stressed that Google is a good tool if one wants to start looking for information in an area of which he or she does not know much. They thought that Google gives a quick overview of the search area and leads to key items or information sources that one can follow for a deeper information-seeking.

Things like Google gives you a sort of a quick and nasty way of getting into something. [AAP1]

- A few (three) of the interviewees, all students, said that they just like the Google brand. These young research students of course have started their academic research as student when Google had already been launched (in 1998) and they might have used it and got used to it ever since the start of their student life. The following quotation is from a student.



I also a kind of like the brand Google because it's very easy and cool and that's why I might use Google first. [SAM6]

- Google is very handy. Google has a simple interface and homepage and nowadays it is part of the web browsers as plug-ins such as Google toolbars are very popular. For example, FireFox Web browser is normally installed with Google search box integrated in it. An interviewee mentioned that Google was actually part of his homepage.
- Google has other functions and applications that could be used by scientists, for example one can conduct calculations or do conversions using Google. It also can be used for equations. These calculation functions were especially popular among students and five of interviewees mentioned use of this application of Google.
- It finds PowerPoint presentations and therefore it has a lot of applications in teaching, and preparation for giving a talk or making a presentation.

When the interviewees were asked to talk about the changes in their information-seeking behaviour over their careers, Google was mentioned by seven interviewees as one of the main sources of changes in their information-seeking habits over time.

The interviewees, however, were also critical in their use of Google in that they were aware of the issues concerning the credibility and accuracy of information. They mentioned that Google presents too many hits for a search and users need to be able to filter through the results to find what suits them.

...I find Google a bit, a bit annoying because no matter what you put in, you get 20,000 answers back. Half of them are referring to the same thing, linked through different ways and you got to be very, very careful what sort of search words you use. You either got too little or too many. [AAP1]

Google is just the world library. The important thing is to be able to discriminate between rubbish, because you know Google will give you a lot of rubbish and things that are not published. [AAA4]

Overall, the study of seeking problem-specific information by physicists and astronomers shows high reliance on use of human and network resources. This is in line with the findings of a recent study by Savolainen (2008) who interviewed 18 individuals active in environmental issues. He found that on average, human and networked sources were favoured in the early phases of information seeking. Printed media such as magazines and organizational sources were often used to complement information received from human sources and the Internet.

## 5.6. Accessing information

As Wilson and Walsh (1995) stated a fundamental requirement for information-seeking is that some source of information should be accessible. The lack of an easily accessible source may inhibit information-seeking altogether, or may impose higher costs than the enquirer is prepared to pay. Once users locate the information they want they need to access the information resource whether it is a journal article, an e-print, a book or any other kind of information resource. Looking for information and making a relevance judgment is normally done based on the surrogates of the information resources i.e. titles, abstracts, subject keywords, excerpts presented in search engines and so on. In this section the issues related to accessing and obtaining information resources are discussed. The information presented here is particularly about accessing and obtaining articles as they appeared to be the main source of information for physicists and astronomers.

Most of journals can be accessed nowadays online as well as in print format which are preserved in libraries. All of the interviewees in the study stated that their preference was to obtain the articles online as PDF files. The reason of course was the convenience associated with accessing an article electronically compared to having to go to the library and reading or photocopying the print version. The participants wanted to be able to access the articles online and if necessary save or/and print them. Statements such as ‘if it is not online I am annoyed’ where the normal kind of expressions the interviewees made in reply to the question about the way they access and obtain the articles. In different subfields of physics and astronomy, the situation is slightly different. Although generally physics and astronomy are among pioneer fields in terms of electronic publishing and online availability of scientific literature (Gould & Pearce 1991), in some fields the situation is better than the others. In astronomy, especially, the online availability is well-advanced.

...for astronomical journals it’s remarkably comprehensive. It really is extremely good. [AAA4]

If an article is not available online, going to the library in order to access the hard copy was the second option that a considerable number of interviewees said they would choose. However, a few (four) of the interviewees stated that before going to the library they would try to see if a colleague has a copy, or they might even contact the author and ask for a copy. Seven of the interviewees thought that if an article is not online then it is not worth the effort to obtain it. They said that if it is not online they would not bother and try to find an alternative source for their wanted information. This surprising statement shows the high uptake of electronic information services among scientist and also changes in the perception of the value of the information sources. It implies the high expectation of scientists for being

able to access all the information they need in the online format. The reason they thought so was the assumption that if an article is a valuable one, even if it is too old to have been published in the electronic format, someone somewhere must have scanned it and put its electronic version on a repository. This is the case for most of classic and important papers in physics and astronomy. The following quotation shows this perception. When the interviewee was asked why he would give up trying to access an article which is not online, he answered:

It's not worth it. So when you know Einstein's papers from 1905, right for instance I don't have to go and get the German right, I know that it's online somewhere because somebody scanned it in, right, so. [AHE4]

In order to see whether this was a common belief among physicists and astronomers, the respondents in the survey were exposed to the following statement and asked to express their level of agreement or disagreement: 'If an article is not available online, it's probably not worth the effort to obtain it.'

The majority of respondents (62.3%) were a little or strongly disagreed with the statements. However, 27.2% of respondents agreed a little or strongly that if a paper is not available online, it is not worth the effort to obtain it. Table 5.5 relates.

**Table 5.5: Distribution of respondents by their level of agreement with the statement: 'If an article is not available online, it's probably not worth the effort to obtain it'.**

	Obtaining non-online articles	
	N	%
I don't know	3	2.6
Strongly agree	10	8.8
Agree a little	21	18.4
Neither agree nor disagree	9	7.9
Disagree a little	30	26.3
Strongly disagree	41	36.0
<b>Total</b>	<b>114</b>	<b>100</b>

Regarding the variation in agreement with the statement among different research groups, 61% of respondents in High Energy Physics (HEP) were a little or strongly agreed with the statement, 17 percent of them were strongly agreed. Astronomy and Astrophysics (AA) was the group with the second highest rate of strong agreement with 14 percent strong agreement. Strong agreement among respondents from HEP and AA groups could be a reflection of high rate of online availability of material in these two fields. These two areas were the first two and probably the most advanced subject areas in arXiv repository. Respondents in Optical Science Laboratory disagreed with the statement. About 90% of respondents from Theoretical Molecular Physics and about 90% of Atmospheric Physics were also a little or strongly disagreed with the statement. See Figure 5.24.

Figure 5.24: Percentage breakdown of responses to the statement on access to non-online articles by respondents' research group.

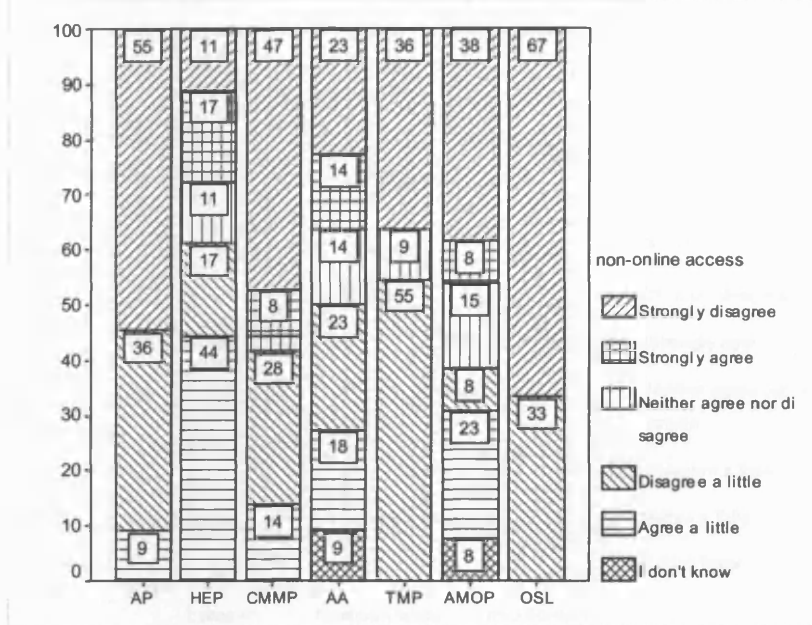
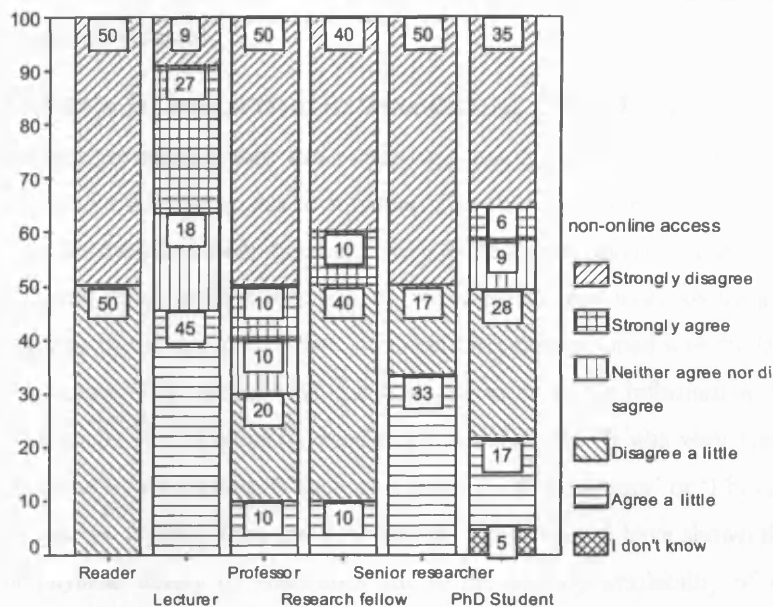


Figure 5.25 gives the percentage breakdown of responses to the statement on access to non-online articles by respondents' status. Lecturers were more likely compared to the others to agree with the statement. 73% of lecturers agreed a little or strongly with the statement, implying that they would not try to obtain a paper if it is not available online. 33% of senior researchers also showed a little agreement with the statement.

Figure 5.25: Percentage breakdown of responses to the statement on access to non-online articles by respondents' status.



This association of online availability with the value of information by some of the participants in the study can have a negative effect as there might be valuable articles that are not available in electronic format and therefore might be neglected. This raises concern about reinvention of the wheel in science, an issue that was mentioned earlier in this chapter (see p. 130).

This approach of users to use of information resources also can be looked at from another perspective, which is perceived accessibility. Users have a perception of the accessibility of the resources they want to use. Past research has shown a strong positive relationship between perceived accessibility and the selection of a particular information source (Gerstberger and Allen, 1968; O'Reilly, 1982; Morrison and Vancouver, 2000). If users think that they may not have access to a particular information resource they would be reluctant to try to use it and would attempt to satisfy their information needs by using another information resource. Accessibility of course has different aspects including cognitive and physical among others. If an information resource is in a language other than the language(s) a user can understand, then that information resource is not accessible to the user even though it is physically at his or her disposal. Ease of access and the principle of least effort plays a part in the choice of information resources and the use of Zipf's least effort principle in providing information services has already been recommended (Bigdeli, 2007). As it was shown earlier the results of the information-event card study show those who used printed journals and

printed books, used books and journals from their own collections and did not go to the library for that purpose. This can be an indication of the effect of accessibility and ease of use on the choice of information resource.

### 5.7. Changes in information-seeking over time

The interviewees were asked if they had noticed any notable difference in their information seeking activities and behaviour over their career. The main outcome of their answers, quite expectedly, was the transition from physical library to the digital library, a shift from the print world to the digital information world. Among interviewees with more senior status and at higher ages such as professors and readers, the main change mentioned was the shift towards electronic information. They all asserted that now almost all of the information they need is electronically available and they hardly need to go to the library. It was very normal for the interviewer to come across statements such as 'I never go to the library' or 'I haven't been to the library for the last 6 years' from the interviewees. Other studies have shown the decrease in the use of physical library by academics due to the desktop availability of information services (Hemminger et al., 2008).

This change towards more electronic information has affected the information-seeking behaviour of the scientists. A high majority (96%) of the interviewees said that they can now conduct their searches faster than they used to do. However the availability of electronic information has been accompanied by considerable increase in the amount of information. The number of publications has increased steadily. Therefore about seventy percent of the interviewees mentioned that although they are faster at their searches, they had to spend more time scanning and filtering information. It has to be said that a few (three) of the interviewees were not sure that their increased efficiency in searching for information was because they have gained experience over their career, or simply because of the advancement in technology.

Certainly a few things have changed a lot. I can now gain the information much quicker than I used to, but this could be also due to the fact that I am more experienced, so the more, some things go together, so I don't know whether I can just entangle the two. [AAA1]

Overall impression from the interviews was that it is a combination of both. As scientists move on in their career, they become well-versed in information-seeking skills and also gain experience in sifting the wheat from the shaft.

The interviewees also talked about their reading. Looking comparatively at what they used to do in the past in the earlier stages of their careers, about 70% of the interviewed staff thought that now they read fewer papers than they used to do. It must be noted that this comparison is not between two scientists with the same status now and in the past, but

between a current senior member of staff and a junior researcher or research students in some years ago. This implies that as scientists progress in their career, they generally read fewer papers i.e. a PhD student or a researcher reads a larger number of papers than a professor does. This is reflected in the following quotation in which the interviewee considers his decreased reading a function of the change in his status.

I probably read less papers than I used to, but that's a function of my current situation. You know you're talking to somebody you know in a very different situation to somebody who will be 20 years younger. [ACM1]

However, firmly reaching such a conclusion requires further research as the interviewees made this comparison regarding their own career and there is a time factor involved here that cannot be ignored. Therefore, reading fewer papers might not be just a matter of status but because the information systems and scholarly communication system have changed. Second, the other aspect is the comparison between the amount of reading an average current scientist does and the amount of reading an average scientists did say 20 years ago. Tenopir and King's (2000) research showed that the number of papers scientists read has increased steadily during the last 30 years. The reason for this might be the increasing number of articles published on every subject. Mabe (2003) showed the steady increase in journal publishing. The ease of access to journals due to the move towards electronic publishing by publishers also might be another reason.

This was also reflected in the interviews, i.e. scientists generally seem to be reading more papers now compared to the past. However, what needs to be looked into critically is that what is considered as 'reading' or how is the process of reading itself. It seems that although scientists might read more papers these days than scientists did 20 years ago, they spent less time reading each paper. In other words, past scientists read the papers more deeply and in more details than modern scientists do. The modern scientists need to scan many papers in order to keep up with the literature of their field. An interviewee stated that with the access to the information the amount of the information exploded, and so that changes scientists' behaviour. While they used to spend more time reading probably journals or searching for journals or ordering journals, there were at the end of the day much fewer papers they would read, but they would read them in more details. He maintained that the way users access the information affects what they read:

...I think sort of the things that I read they are prescribed by the way we access the information right so, because I realise that, but there is very little we can do. Because when you choose what to read you know the choice I think is sort of given to you now by the fact that you are in this particular field and these people read this particular number of magazines and you don't read the others and you realise that but you can't do everything. You have to restrict yourself somehow to these activities....I spend more time in reflecting, sort of analysing the information than

for looking. I used to spend more time in looking for information before, but now it's really very accessible, but also the number of papers you have to sort of look through is also much bigger. [ACM8]

A professor believed that this transition period from print to the electronic information environment has reached a stable era and it is already over.

It is not really changing. The journals have only come online in the last ten years so there has been ...because when we were having to get that whole process through. It is now it got to a stable state where you submit online, you do everything online. We had to go through intermediate status where you post it at the end and they posted it back, then we got them to email things. It is sort of, it has reached now I think a stable situation whereas 15 years ago everything was paper. So it has been this transition. I think the transition is just about over now. [AAA5]

The other significant driving factor of changes in the information-seeking behaviour of the participants was Google. Google has affected the information-seeking behaviour of those who use the Web considerably. A review article by Bawden and Vilar (2006) shows that use of Google has changed the perception of users, especially younger generations, of online resources and digital libraries (See page 157 for further discussion on the role of Google in information seeking of the participants.)

...as soon Google came that made life easier as well. I don't think I go about 20 minutes without using Google. Because even within our experiment to actually find something within all the documentation, my own experiment, I have to use Google to find the information on my own experiment. [AHE4]

For those who were in early stages of their career such as young researchers, the main change was being more independent and relying on their own skills and judgement for information seeking. During a PhD course, students rely a great deal on their supervisors and colleagues for finding information, but as a research fellow, one is more independent. A researcher described this change as 'seeking his own information':

I guess I'm much more independent now. Like at the beginning like I would just read the papers that my supervisor told me to read. Pretty much, so every time I needed to know something I maybe go to him and he would say read this paper by such and such, but now as I've obviously gone on I learnt the literature more and I learnt the people more. So now I find... I look for something I think this is something that such and such would do, so I search for the person and now I seek my own information. Whereas before like I always relied on others. [AAA7]

Although all of the interviewees had a general sense of satisfaction with the move toward electronic information and the ease and convenience associated with it, this change has come with some disadvantages at least for some senior interviewees who had experienced the older print information environment. Six interviewees among staff had complaints. The main disadvantage complained about was the practical difficulty of browsing. Browsing is a very convenient and joyful act when it is done with print journals. But in the electronic



environment it is neither joyful nor convenient as interviewees believed (This was further discussed under section 'Browsing and serendipity' in page 204).

## 5.8. Reading behaviour

In the interviews, participants were asked about their reading behaviour. The questionnaire survey also included questions about the number of preprints and articles they read during an average month, the age of the last article they read and the method they used to find it. This section presents the findings of the interviews and questionnaire survey on the reading behaviour of the participants.

### 5.8.1. Reading quantity

Respondents of the questionnaire survey were asked about the number of articles and preprints they read. It appeared that the respondents read more articles than preprints (Table 5.6). The median for the number of articles read was six while for preprint was three; the mean values were about eight and five respectively.

Table 5.6: Number of articles and preprints read.

No of articles or preprint read	Article		Preprint	
	No of participants	%	No of participants	%
<= 5	51	44.7	83	72.8
6 - 10	44	38.6	20	17.5
11 - 15	6	5.3	5	4.4
16+	13	11.4	6	5.3
Total	114	100	114	100

### 5.8.2. Reading and other activities

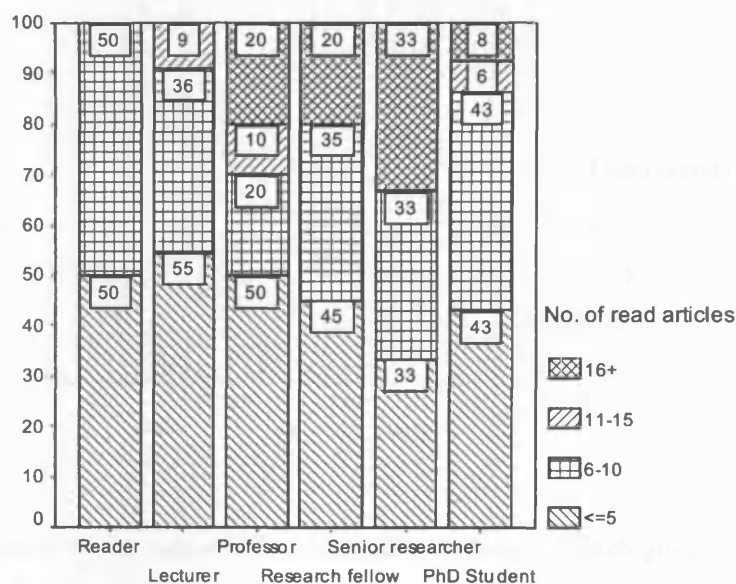
Pearson correlation coefficient ( $r$ ) for the number of articles read and the number of articles published was computed 0.109, which is an indication of a positive correlation between these two variables, even though a very weak correlation. It means that those who published more articles tended to read more articles.  $r$  was 0.235 for preprint indicating an slightly stronger positive correlation between the number of published articles and the number of preprints read. To find out about any correlation between the amount of reading and the time spent teaching or writing (articles, grant proposals and etc.). Pearson correlation coefficient was computed. While there was a positive correlation between writing and reading ( $r = 0.188$ ), the correlation was negative for teaching ( $r = -0.045$ ). This shows that those who spent a higher proportion of their time teaching tended to read fewer articles, and those who spent more time writing tended to read more articles as well. It should be noted that all these correlations are very

weak according to these figures and they merit further investigation perhaps by surveying a bigger population. However, past research has also shown a positive correlation between amount of reading and the scientific productivity. For example a survey by Tenopir and King (1996) showed that scientists who read more are more productive and perform their work better. Tenopir and King (2002) also found that a correlation between reading journal articles and professional achievement - award winners read more articles than non-winners.

This result is also illustrated in Figure 5.26 as those who are allegedly more involved in research and writing (e.g. researchers) are more likely to read a larger number of journal articles. A third of senior researchers, a fifth of research fellows, and a fifth of professors said they read 16 or more articles a month while this figure for lecturers was zero.

Table 5.7 also shows the number of articles and preprints read by academic status and it shows that lecturers read the least number of articles.

**Figure 5.26: Percentage frequency distribution of respondents by status by number of articles read.**

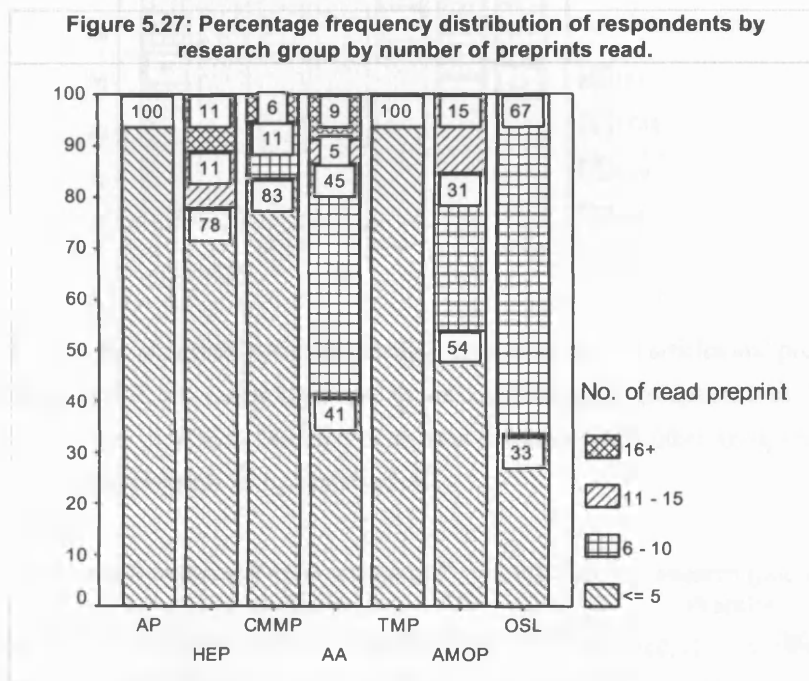


**Table 5.7: Number of articles and preprints read by academic status.**

	Article		Preprint	
	Mean	Median	Mean	Median
Student	7.6	6	4.7	3
Research fellow	9.2	9.2	4.1	2
Senior researcher	11	10	3.5	2.5
Lecturer	5.4	5	8	5
Reader	7.5	7.5	3.5	3.5
Professor	10.2	5.5	7.7	5

### 5.8.3. By research group

Not surprisingly there were differences in the number of preprints read among different research groups (Figure 5.27). People in Astronomy and Astrophysics read the highest number of preprints with 59% of them reading more than 6 preprints a month. High Energy Physics research group included the biggest number of people (11%) who read 16 or more preprints a month. None of the respondents in Atmospheric Physics or Theoretical Molecular Physics read more than six preprints a month. Those who read 16 or more preprints a month tended to be from High Energy Physics (11%), Condensed Matter and Material Physics (CMMP) (6%) and Astronomy and Astrophysics (9%).



With regard to the number of articles read by different research groups (Figure 5.28), Atmospheric Physics had the highest rate of reading articles with a fourth of them read 16 or more articles a month. They were followed by respondents from AMOP group from whom 15% read 16 or more articles a month. On the hand, High Energy Physics people had the lowest amount of article reading with the majority of them (83%) read five or fewer articles in a month and the remaining 17% read 6-10 articles. None of the respondents from this group read 16 or more articles in a month. This makes sense as this group had the highest rate of preprint reading which means they rely more on preprints rather than articles. Generally those

who rely more on journal articles read fewer preprints and those who rely more on preprints read fewer journal articles.

Figure 5.28: Percentage frequency distribution of respondents by research group by number of articles read.

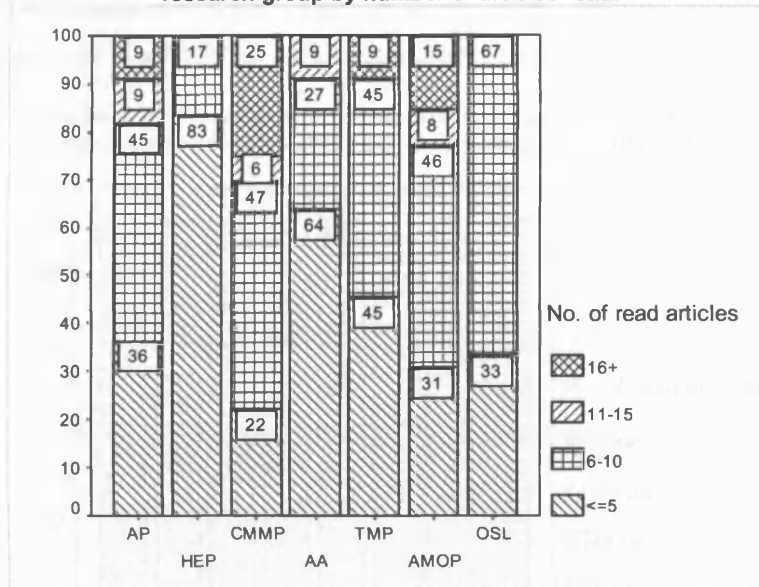


Table 5.8 also shows the mean and median number of journal articles and preprints read by respondents in different research groups. As we could see some groups such as CMMP, AP and TMP read more published journal articles than preprints while other groups such as AA and HEP read more preprints than journal articles.

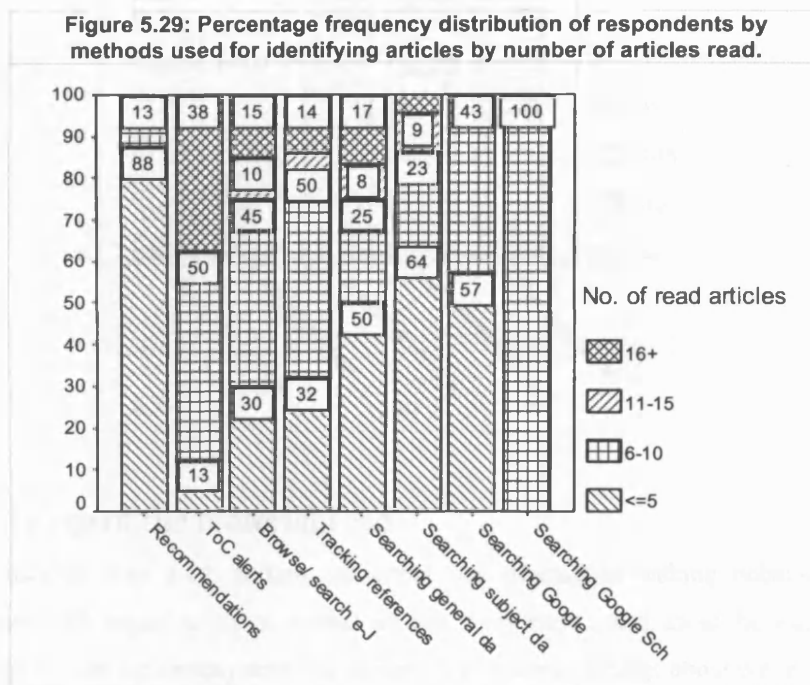
Table 5.8: Number of journal articles and preprints read by research group.

	Article		Preprint	
	Mean	Median	Mean	Median
AA	5.9	5	8	7
AMOP	9.8	8	5.9	5
AP	7.3	6	1.6	1
CMMP	11.6	10	4.3	2
HEP	3.2	2	5.9	3
OSL	8.3	10	8	10
TMP	7.5	6	2.1	1

#### 5.8.4. By methods used for identifying articles

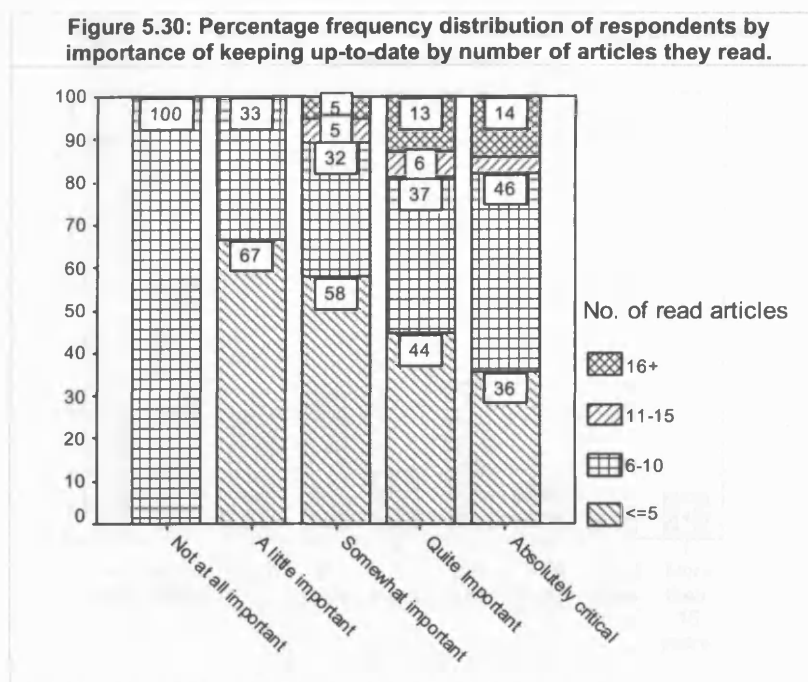
Figure 5.29 demonstrates the correlation between respondents' favourite methods for finding articles and the number of articles they read. Those who read a larger number of articles in a month tended to rely on ToC email alerts, searching general databases, browsing e-journals and tracking references more than the others did. Those who used recommendations by

colleagues as their main method for finding articles were most likely to read the fewest number of articles compared to the other respondents. Those who used ToC email alerts for finding research articles were more likely to read more articles than the others; 38% of them read 16 articles or more in a month. Those who depended on colleagues' recommendation for finding articles were least likely to read a large number of articles.



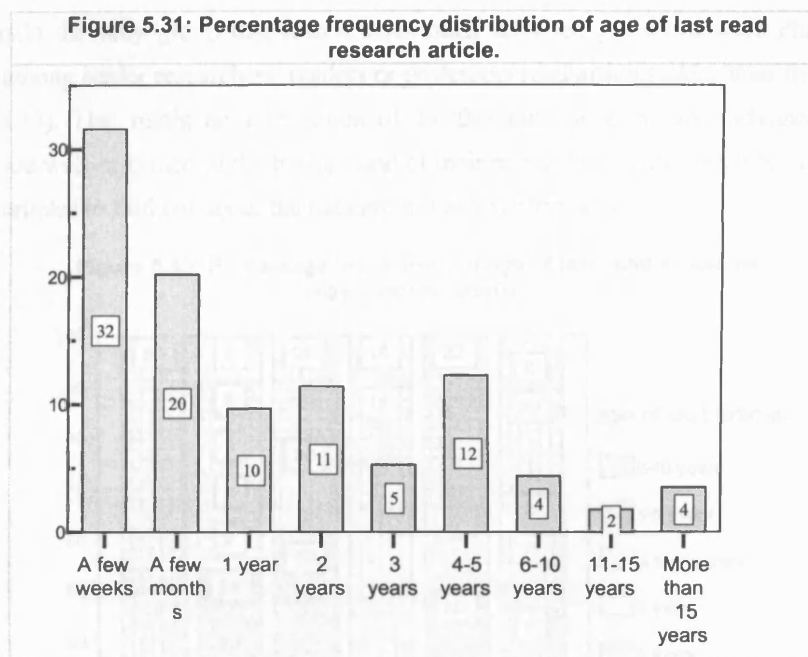
### 5.8.5. By importance of keeping up-to-date

Those who associated more importance with keeping up-to-date were more likely to read more articles. Eleven percent of those who said keeping up-to-date was absolutely critical and 13% of those who said it was quite important read 16 or more articles a month. Those who considered keeping up-to-date a little important did not read more than 10 articles a month. The only respondent who said that keeping up-to-date was not important at all for him read 6-10 articles a month. See Figure 5.30.

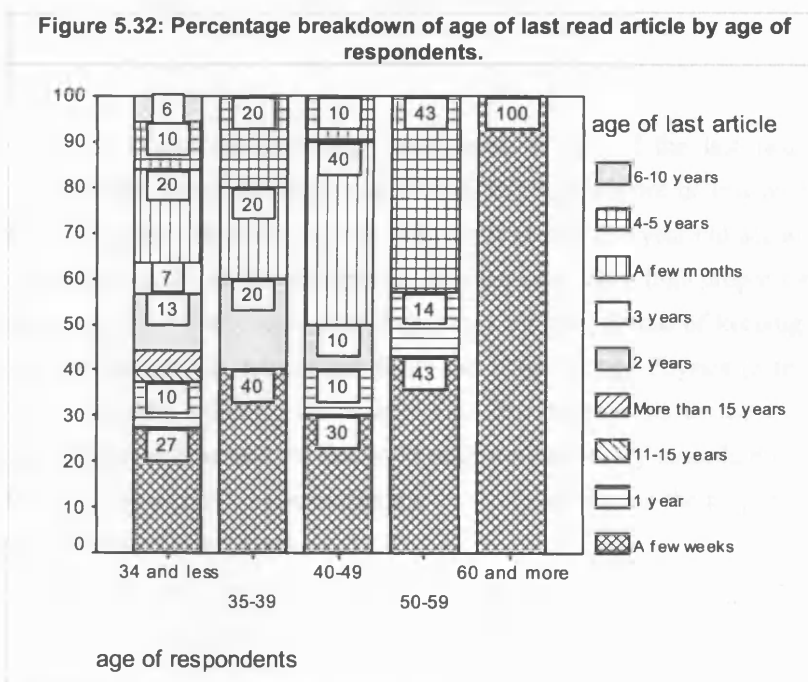


### 5.8.6. By age of the material read

Critical incident was used to find out about the information-seeking behaviour of the respondents with regard to use of journal articles. They were asked about the way by which they found the last article they read and its age. The detailed findings about the methods used for finding articles were presented in the section on Identifying Articles (page 120). Regarding the age of the last article read (Figure 5.31), about a third (32%) of them were a few weeks old, which is an indication of reading articles for keeping up-to-date probably. Six percent of the articles were more than ten years old. These results resemble the findings of Tenopir and colleagues' (2003) longitudinal study. They showed that the distributions of age of articles read were very similar over time since early 1990s. For example during the two periods of 1990-1993 and 2000-2002, respectively 65% and 69% of the articles read by scientists surveyed were maximum one year old, and eight and five percent of the articles were more than ten years old. In the current study, 62% of the articles read were one year or less old, and six percent of the articles read were more than ten years old. The difference between these figures and Tenopir's figures is not considerable.



The only age group of respondents whose last read article were six or more years old was those 34 or fewer years old (Figure 5.32). Respondents of age range 60 or more all had read a few week old articles.



Again the only group that read articles older than ten years old were PhD students. Nobody among senior researchers, readers or professors read articles older than five years old (Figure 5.33). This might be a reflection of the fact that those in more advanced level of research are well-informed of the background of their research area and less often may need to read old articles to find out about the background of a subject area.

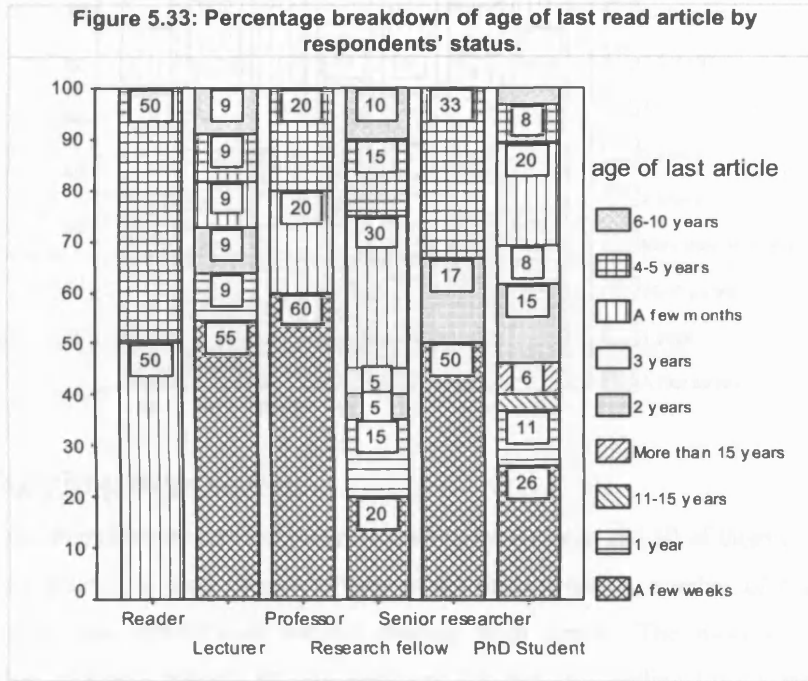
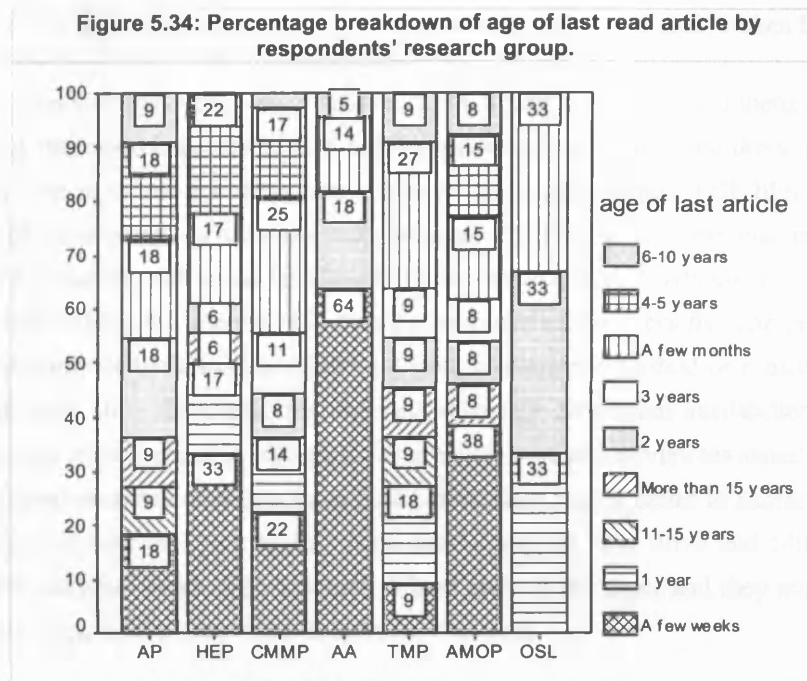


Figure 5.34 shows the percentage breakdown of age of the last read article by respondents' research group. Older articles seemed to be used more or less by respondents from all research groups. However, articles from a few weeks to a year old accounted for the majority of articles read by all research groups. This indicates that a high proportion of articles read by respondents are from recent articles probably for the purpose of keeping up-to-date. Astronomy and Astrophysics (about two-third) and High Energy Physics (a third) had the highest rate of reading articles a few weeks old. The respondents in the two groups of Atmospheric Physics and Atomic Molecular Physics were relatively more likely to read older articles. These two groups were also among those who relied less on the preprints and tended to read more journal articles instead.





### 5.8.7. Reading from screen

None of the interviewees were in favour of reading from screen and all of them clearly stated that if they wanted to read an article they would print it out. A number of reasons were mentioned by the interviewees for not reading from screen. The main reason was a combination of health concern for eye sight and the fact that reading from screen is not a convenient thing to do. The following quotation from a student is very indicative:

Although LCDs are much more comfortable these days but having a paper on your hand, something you can scribble on directly, it's much more comfortable. This reminds me of the story of television you see, when television came around people were saying radio is going to die, it's going to disappear. People would require paper, paper is a much more comfortable media to read and to manipulate. There was a discussion about this and someone was saying its' just a question of getting used to it, if you get used to reading on screen there is no need for paper. No, there is nothing like turning a paper, the way you manipulate the paper, you paper is more humane let's put it this way. [SAM1]

However, users do not print every single paper that they retrieve. The interviewees did some screen-reading and the amount of screen-reading varies from one person to another. Most of the interviewees stated that they would print out the paper if they want to read it thoroughly or if they want to take the hard copy and read while commuting for example. But when they look for specific pieces of information in a paper or they just want to have a quick look at it they would not print it. There was also concern over wasting paper and a few of the interviewees mentioned that they try to print fewer articles in order to avoid wasting paper.

If I only want to look quickly I will do it on screen. If I want to study it then I do it on paper. It's much easier on paper. [ACM10]

A researcher may print tens of articles during a year in order to read them, and also he or she may download hundreds of PDF articles and save them on the hard drive of his or her computer. The questions that raises here is how do users manage these PDF files and printed articles, or more generally their personal information? This is an issue that merits to be researched separately and could be subject of a dissertation itself<sup>16</sup>. Although this issue was not discussed during the interviews in detail, it was touched upon briefly. The outcome was that the majority of the interviewees did not have a systematic method or a mechanism for managing these files. Of course, the easy and relatively convenient availability of journal articles online playing a part in this issue. For example, several interviewees stated that if they need to reread an article that they have read before, they find it easier to access the article online again instead of trying to find it on their computer hard drive and filing cabinet. Sometimes, they just forget that they have a hard copy of the paper and they might end up printing the same article a few times in different occasions.

It has to be said that twelve of the interviewees (mostly students, nine out of twelve) used reference management software and tried to organise their notes and files. However, some believed that the number of articles published on every subject is just too big to try to keep a personal archive of the relevant articles.

### 5.8.8. Role of abstracts

Articles are found and identified by users using different methods. Some methods such as recommendations by colleagues and chasing references at the end of articles are very selective in nature and will provide the searchers with just a few papers. However, some other methods such as a search in a database might result in numerous articles, a list that users need to go through and decide which ones are relevant and should be read. The first parts of an article that a user face is its title and abstract that act as surrogates of the full-text. Abstracts, despite having a long history, are still the subject of much comment, debate, and research (Nicholas, Huntington and Jamali, 2007). However, abstracts still seem to play an important role in the relevance judgment that users make when they browse a list of articles. Pinto and Lancaster (1999) pointed out that, despite the changing role of abstracts from 'a human-readable output from electronic databases', to 'computer-searchable surrogates for larger bodies of text' and with the arrival of ubiquitous full-text 'They remain useful summaries to be read by humans'

---

<sup>16</sup> As a matter of fact as the importance of this issue has grown the British Library has started a research project named 'Virtual Lives' that investigates this area. For more information about the project see <http://www.bl.uk/digital-lives/>

and protect the searcher from ‘unacceptable levels of irrelevancy’ that result from full-text searching. One of interesting findings of the Virtual Scholar Research programme<sup>17</sup> run by CIBER is how popular abstracts proved to be. This is interesting because the ease of full-text availability has made it easy to consult the full-text of articles and scan through them in order to decide about their relevance, something that one of the interviewees did. A research fellow [ATM3] said that he normally scans quickly through the articles and has a look at their introduction and conclusion to decide whether he wants to read them fully or not.

However, few people said they would do that and the reliance on abstracts was high among the interviewees. Forty nine (about 87%) of them said they would check the title first and if the title was interesting enough then they would read the abstract and if the abstract seemed relevant they would download the article and skim through the article itself.

Although abstracts were perceived to be useful, there were three interviewees who complained about ambiguity of some abstracts that do not present the content of articles properly or sometimes even mispresent the content. This results in waste of time for readers.

A relevant point that must be said here is that some interviewees also use the names of authors as a relevance criteria so if they spotted a paper by someone whom they know is a key author on the subject, they would be more inclined to download that article and read it.

A professor developed a systematic mechanism for making the most of abstracts. He said that he receives email alerts in his areas of interest, and then he copies the interesting abstracts on a subject to a MS Word file with the same name. He has about 40-50 different topic areas with those abstracts in. Every time that he wants to check one of his favourite subjects for a research he would check his abstract collection. He described the email alert service as follows:

One of the areas I work on is conducting in semi-conducting polymers. When I pick up an abstract of interest from one of my searches I simply copy it into a word file here under the heading ‘organics’. So this has got 124 pages. That’s rather a lot of them and... And so what I have at home is about 40 or 50 different topic areas with those abstracts in. Now the first place I would go if I wanted a particular topic might be the relevant abstract list and just look what’s there. [ACM3]

### 5.8.9. Credibility of information

Information found on the internet can sometimes be problematic. It is always important for users to evaluate information found. Credibility can be simply defined as ‘believability’ (Fogg, 1999; Tseng & Fogg, 1999). Credible sources are described as ‘trustworthy’ and having ‘expertise’ (Self, 1996). Of course if the information found in a reliable source such as a

<sup>17</sup> For information about the programme see: <http://www.ucl.ac.uk/slais/research/ciber/>

journal article that has gone through a reasonable quality control process such as peer review it will be more reliable. However, nowadays there are a lot of preprints deposited in repositories as well as other types of web-based information sources that could be used by scholars. Some of these resources do not go through any type of quality control. For example preprints are not peer reviewed. The credibility of this information and the criteria that scientists apply for evaluating this information is an important issue. While discussing reading issues with interviewees, evaluation of information was also discussed and interviewees were asked about their method of evaluating the credibility of information if they did.

Most of the interviewed students were aware of the issues concerning credibility and validity of information resources, particularly journal articles and preprints. As mentioned earlier, published journal articles are trusted due to the fact that they are peer reviewed, preprints are treated with caution. Wathen and Burkell (2002:141) proposed a model of how credibility assessment may occur on-line. According to their model one of the questions that users may or should ask themselves for assessment is 'am I ready to act on this information'. The current study showed that the interviewees do ask this question and adopt different levels of rigorousness in their assessment of the credibility of information. For example the interviews showed that those who read preprints read them first of all for keeping up-to-date. If the study presented in a preprint is to have some impact on their research, in other words if they are to act on the information found in a preprint or a Web resource students tend to discuss it in detail with their supervisors and colleagues and if possible check the results. They try to obtain the published version of preprints if they sound interesting to them. Past studies (Budd and Connaway, 1997; Speier et al., 1999) also showed that the perception of quality of information sources (which is a relevant attribute to what we discuss here) is an important aspect in whether or not. According to a student:

'in physics, things are typically right or wrong, there isn't much scope for interpretation or 'opinion' [SHE6].

In collaborative experimental physics, particularly high energy physics, papers are more likely to be trusted easily because they are produced by large numbers of authors (sometimes by hundreds) and they go through strict screening processes before a collaboration releases them to the public. In these areas of physics, the names of authors tend to have less significance in the readers' judgement on papers' credibility simply because there are too many authors on the papers, and it is not easy to remember the names. Meanwhile, in theoretical physics, names of authors and their reputation are more likely to be considered as a factor in the judgement.

An interviewee stated that in order to see whether they can rely on a paper, they 'do not read just one paper but look at the spectrum of papers [SCM1]'. In highly competitive subfields of physics, there is a sense of trust on collective knowledge because they believe if some faulty results get published, they would get found very quickly by the subject community. A PhD student stated that for him one way of evaluating a preprint that has not been peer reviewed was to see who the author(s) was. He maintained that due to attendance at numerous conferences he had a good knowledge of the people in his field. However, he also believed that 'good people write bad articles' too; therefore he would never really rely on an article without checking its results.

I read the beginning of the article and see if it makes any sense to me. Okay for adding a citation to my own publication I don't have to check the whole article. If I wan to cite what has been done in the field yet then I can add my citations without having checked all the results; though I prefer to cite published papers of course. If I want to rely on the results by using them and building something new based upon those results then I would check it. [SAM2]

In the information-event card study only one user made a comment on one of the cards regarding this issue. A student [SCM4] used some web-based information resources such as webelements.com and Wikipedia to find the melting and boiling points of potassium. During his search for information he realised that there were some variations in quoted values and therefore he felt the need for cross-checking to make sure about the accuracy of the values. He eventually decided to use the values that were mentioned in the majority of the resources he consulted [SCM4]. This is an interesting case to show how critical the validity of online information can be and how critical users need to be in using the information, something that two recent reviews of credibility (Wathen and Burkell, 2002; Metzger, 2007) highlighted as well.

### **5.9. Publishing behaviour**

Arguably publishing is the main means by which scholars communicate their research outcome. In scientific fields such as physical sciences, journal articles are the main form of publication (Tenopir and King, 2000). To write and publish an article, scientists go through a process which involves some information seeking activities. Interviewees discussed this process during the interviews. It is worth mentioning that there are different types of articles including for example review articles, research articles and so on. What was discussed by the interviewees was research articles that present the findings of their research studies. It should be mentioned that PhD students were also asked about publishing. However, most of them had little experience in terms of publishing and all stated that the decisions related to publishing are made by their supervisors or at least with their advice.

### 5.9.1. Writing approach

Half-jokingly and half seriously, many of the interviewees described the process of writing a paper as a 'painful' process. There were variations in the interviewees' approaches to writing. For example a theoretician professor in high energy physics [AHE3] said that he would write the paper entirely in his head before actually starting the writing. He would have everything laid out in his head so the actual act of writing would be fairly easy. While some authors [e.g. AAP2] start the writing by writing a brief abstract and brief introduction and conclusion and the result section and then try to expand each section, some others [e.g AAA3] would write the results section first; and the introduction and conclusion would be the last sections they would attempt to write.

However, regardless of all these variations in the order of writing, there seemed to be a common process for writing a paper. The writing process starts when there are some results that scientists think they are worth publishing, something like a new model, a new theory, a new experiment or new data. At this stage, most of the interviewees said they would write the bulk of the paper which presents the new data and findings. Then they would try to write the introduction and the conclusion in order to explain what and why they try to say in the paper. At this stage they would need to put their findings and data in context by citing past studies and linking their data to some background data. This stage involves some information seeking activities and literature searching.

As stated before one of the two main searches for the literature that the interviewees mentioned they would do during the course of a project is close to the end of a project or at the time of writing. This search is less thorough and more focused compared to the literature search conducted at the beginning of the project for two reasons. Firstly, a researcher already would know most of the relevant literature because of the literature search that has been conducted at the beginning of the project. Secondly, researchers would try to keep abreast with the relevant literature during the course of a project and therefore they would be aware of any recent paper or development that has emerged since the start of their research.

The aim of this literature search at the time of writing, as one of the interviewees stated, is:

A) You know everyone else's past work and B) you want to make sure that you reference everybody that you should be referencing. You don't want to miss any person. [AAA1]

The nature of the search for and use of literature at the time of writing is also different from the literature search at the beginning. Most of the interviewees said that they would refer to the papers that they already know and try to use and reference them in order to put their

own research in context. They would conduct searches in order to fill the gaps they might encounter in their writing. The search at this stage is also more focused and specific because researchers perfectly know their research and results and as one of the interviewee [AAA4] said at this stage they know exactly what they are looking for, if they search.

As a professor (quotation below) explained, the search for the literature would be just for recent literature, roughly a year or less old. This is because they consider themselves as the experts in that area and reasonably they expect to know about the older literature in that area. The other reason is that they have already done a more comprehensive literature search at the time of writing the grant proposals or the research proposals for the funding body and they use that for their writings. The literature search at the time of writing is to make sure that they do not miss any recent significant papers on the subject.

You will do that in terms of tying in your new results with perhaps somebody else's new results, fairly recent, but anything more than a year old let us say you will know it, because you are an expert in the field. You will already know that, so I won't do a completely new literature search going back more than about a year for a new paper, because it is not really relevant. That is already covered in the previous paper so it is a kind of linear progression in terms of publishing things.  
[AAA3]

### 5.9.2. Publishing approach, journals

An important decision to be made by authors after or before writing a paper is the decision about the target journal as an author's choice of a journal can affect whether or not his or her paper is published and how quickly it is published as well as the impacts it can have on his or her professional reward and social recognition. Past studies (e.g. Gordon, 1984; Frank, 1994) have investigated several factors that scientists may take into account for deciding about the journal in which they want to publish their articles. In the field of physics a survey by CIBER (Nicholas et al, 2005b) investigated this issue. They asked physicists to say what their main reason(s) was for publishing in the journal in which they last published. Respondents were offered 22 possible reasons with the option of saying if this was a reason (Yes) or not a reason (No). For many authors choice was a complicated and personal matter as evidenced by the comments the authors left. The main reason for selecting the last journal in which authors published was that that journal covered their area of interest: just under two-thirds of respondents said this. Other important factors, albeit some way behind, were prompt publication (14%), worldwide readership (14%) and high impact factor (13%). A survey of 683 researchers in astrophysics and cosmology by Polydoratou and Moyle (2006) showed that the majority of them (494 people) considered the quality of the journal as perceived by the scientific community a very important factor affecting their decision for selecting journals.

The findings of interviews in the current study showed variations amongst subfields of physics and astronomy in terms of the top factors they consider at the time of making this decision.

In astronomy generally the journal choices are few because there are just three journals that are considered by researchers (in a prestigious research group such as the UCL Department) good enough to publish in. As a professor listed, they are:

In astronomy there are apart from Nature, there are three main journals. One is American, *Astrophysical Journal*, one is European, called *Astronomy and Astrophysics*, and one is the UK journal called *Monthly Notices*. [AAA3]

It appeared that the cost of publishing was the main driving factors in deciding which journal to submit to in astronomy and astrophysics. The American journals tend to be considered more prestigious, and many people want to publish in those journals. But a decisive factor here is that the American journal (*Astrophysical Journal*) has page charges that authors have to pay. British research funders in astronomy and astrophysics (mainly the Engineering & Physical Sciences Research Council, EPSRC) do not seem to like this policy and therefore British authors a lot of time publish in the British or European journals in astronomy if they cannot cover the publishing charges. When asked about how he would make the decision as to which journal to submit to, a professor in astronomy answered:

Cost; There are three leading journals in astronomy and one of them is American and has page charges and then the other two, one is European and one is British. Although the British one is an international journal it just happens to be based in Britain and I submit to the British one. They're both without page charges but the publication standards of the British one are higher generally... [AAA4]

He also maintained that many British people would probably publish in the American journal if they had the funding, and when he was asked why they would do so he answered:

Because it's perceived as being read by more astronomers. I don't think that's true and I think with the way that electronic is going how it's perceived is changing. Even if it was true at one time it's less true now. [AAA4]

This cost factor was mentioned by all of the astronomers and astrophysicists interviewed. *The Astrophysical Journal* which is an American journal has page charges, therefore one needs to have enough money to pay for page charges if wants to publish in the American journals. A professor [AAA5] in astronomy explained that the astronomical journals outside the USA do not charge. The *Astrophysical Journal* is probably the most prestigious of the pure astronomy journals and they do not waive page charges. Other journals like *Monthly Notices of the Royal Astronomical Society* and *Astronomy and Astrophysics*, which is the European journal, have no page charges. Therefore, here the issue is the scope and orientation of the journal. *Monthly Notices* are more oriental extra galactic and cosmology



whereas *Astronomy and Astrophysics* is more galactic astronomy, stars and so on because it reflects the European research interest on the continent. But there may be other criteria, like *Monthly Notices* has copy editors at Blackwell Publishing who correct the English, so if one is a Chinese author who cannot write good English, they will actually correct the English into proper English, without mangling the science. Where *Astronomy and Astrophysics* tends to take the papers, virtually as abstract and camera-ready form and publish it without that kind of corrections therefore some papers come out in 'pigeon English'.

These few options and important differences they have in terms of the author-charge, has made the decision-making process very easy in astronomy and astrophysics. Several interviewees said that this is a default decision in astronomy. Many people tend to keep publishing in the same journal. Sometimes other factors affect the decision about the journal. For example whether the co-authors of the paper and the project collaborators are American or European might make the American or the European journals more favoured.

In physics, the journals are more numerous and authors have a wider range of possibilities for deciding which journal to publish in. This also makes this decision-making process more complicated. Because when there are more journals with different qualities, frequencies, editorial boards, charge-policies and so on, authors have to take more factors and elements into account. For example a professor in atmospheric physics in reply to a question about the factors he considers for making this decision answered that there is a combination of factors including visibility, importance and speed. Referring to a short and 'good' paper of his, he went on

...we are trying to work out what to do with this one. It is a fairly short one and we think quite a good idea. We want to get it out quickly. Now if you know if you send it to *Annals de Physique* which is a European one, and you could get it produced in [*Physical Review*] Letters, then they will produce it fairly quickly. The impact of that journal is not very high, so if you send it to one of the American journals where the impact is higher, you could wait ages for it to be produced. So you have to decide which one you think is the most important, do you want this to be a high impact paper, or do you want this to be one that you get out quickly... It depends how quickly you want it out and how much impact you want it to have, how much, what you think it is worth. I mean some papers you sort of produce and think you don't want to send those off to the most important journals, because they are not important enough to warrant it, but others you may think are a really important result. You want to maximise the outcome. Nature is a classic example. If you got something you think is absolutely a unique result, and it is newsworthy, then you send it to Nature. [AAP1]

However, this does not mean that factors such as speed and quality overshadow the relevance of the journal. The factors such as quality and speed are secondary to the relevance of a paper to a journal's coverage and audience, something that CIBER's survey also concluded (Nicholas et al., 2005b). This is an issue that is more noticeable in subfields such as

Condensed Matter and Material Physics (CMMP) that are more interdisciplinary and have the possibility of publishing in even a wider range of journals. A professor [ACM1] from CMMP group stated this clearly. He stated that it depends on the audience and the nature of the work. Sometimes the kind of things that the journals publish is on tradition. For example a physicist might publish in a chemistry journal if the interest of the work is largely to the chemist even though the authors are physicist. They might be giving a physical explanation that for instance a solution has a solvent property they have been observing in a specific liquid. Of course an author must be able to convince the editors that the paper is of importance to that community. The professor also stated that

There's largely a community, I mean you're looking at the audience that you want to read the paper. I tend to go for American journals largely because the Americans don't read the non-American literature. That can be very parochial in not looking at what's published outside the US. So, yeah. [ACM1]

A researcher in the Theoretical Molecular Physics [ATM2] had his own way of deciding about the relevance of his paper with a target journal and its audience. He stated that he would look at the references of his paper and would submit to the journal from which more articles have been cited in his paper. The rationale for this is that more citations from a specific journal means the article is more likely to be relevant to that journal and be read by its audience. However, he also mentioned that he would also look at the quality of his paper and decide to publish in the best possible journal amongst the relevant ones. There are other factors that might have bearing in this decision. For example an editor of a journal in CMMP said that he would normally publish in his own journal.

In a normal situation when the cost issue is not a barrier, when there are a few relevant journals to choose from and when the author knows how quickly he or she wants his/her paper to be out, the decision is to publish in the best journal possible. This is of course after the assessment of the paper's quality and considering the chance the paper may have for getting published in the given journal. Highly prestigious journals have a higher rate of article rejection so authors have to balance the risk of getting rejected against if the paper gets accepted it is much more prestigious. A professor explained that:

You always try the best journal in the field. Most journals are now specialised. The world's biggest physics journal is called The Physical Review. It's an American journal and comes in about eight different sections, particle physics, condensed matter and so on. So you choose the journal which is in the right field. And you choose the journal, which in the jargon, has high impact factor, which means there are statistics which say that people read that journal. If you publish in an obscure journal, the chance is no one is going to read it. You can always find somewhere to publish your article, but the fact is every subject has a hierarchy of journals. You do that because then people read your paper, so it's prestigious for you. [AHE3]

While in astronomy deciding about journal was a decision that the interviewees considered as something known by 'default' and clear before the paper is written or finished, in physics this decision might be made once the paper is done. The reason is that sometimes the authors need to assess their final draft of the paper in terms of quality and then decide where to send it.

### 5.9.3. Publishing approach, e-prints

Statistics of e-print repositories show the steady growth in their usage as well as in the paper submission rate (arXiv, 2007). Physicists and astronomers are among pioneering scientists in terms of using repositories. They have had a long tradition of circulating preprints of their articles among their colleagues and they were the community who established the first e-print archive. Los Alamos National Laboratory e-print archive was established by Paul Ginsparg, a physicist, in 1991 (Ginsparg, 1994).

The use of and the interaction with e-print archives were discussed during the interviews and also a few questions about use of e-print archives were included in the questionnaire survey. Table 5.9 shows that the respondents of the questionnaire survey were distributed relatively evenly with regard to their habit as to whether deposit their articles in e-print archives or not. 37% did and 38% did not. 25% also said that they did not know about depositing; they were all PhD students in early years of their studies that they did not know about depositing or had not written any articles. From those who said they did deposit their articles, 36% did it before they submitted their articles to journals, 33% said they did it at the time of submission and 31% did it after their articles were accepted (Table 5.10).

**Table 5.9: Distribution of respondents by whether they deposit their articles in e-print archives.**

Depositing	N	%
Don't know	29	25
No	43	38
Yes	42	37
Total	114	100

**Table 5.10: Time of depositing articles in e-print archives.**

Time of depositing	N	%
Before I submit	15	36
When I submit	14	33
Once accepted	13	31
Total	42	100

As Table 5.11 illustrates, the most common reason for not depositing articles in e-print archives was that they thought that it was not common or a tradition in the subfields of those who did not deposit. 44% of those who said they did not deposit opted for this reason. The other common reason was the statement 'I can't be bothered', with 21% of those respondents who did not deposit. A few of respondents used the provided textbox to mention some other reasons. Two respondents were concerned about their ideas being stolen. One respondent wrote 'sometimes it's better to keep some results secret for 3 months longer'. Another one said 'not been important so far - use ISIS facility (best place in world for neutron diffraction from liquids) which means that no one's going to repeat my experiment!' One respondent also blamed lack of time. In the interviews, some respondents stated that they were simply being lazy or they could not be bothered to deposit their papers. It is also worth mentioning that most of the papers in physics and astronomy have multiple authors. In those situations when there is a student or a younger research fellow among the authors, they are the ones who do the depositing and older authors such as professors are less likely to do it personally. This is contradicting the findings of Wertman (1999) who interviewed 12 physicists and concluded that mid-level and older scientists were consulting the e-print archives more than the younger scientists. This may not be surprising as Wertman's research is already almost a decade old and things have changed since then.

**Table 5.11: Reasons why not depositing articles in e-print archives.**

The reason	N	%
Doesn't have any benefits	5	12
It's not common	19	44
Copyright doesn't allow	1	2
Can't be bothered	9	21
Other	9	21
Total	43	100

As we can see in Figure 5.35, 64% of respondents in Astronomy and Astrophysics deposited their articles, which is the highest figure among all research groups. People in Atomic, Molecular, Optical and Positron Physics (AMOP) research group (54%) and High Energy Physics (44%) also had a high rate of depositing articles. In some areas such as HEP, physicists nowadays must deposit the preprint of their papers in the archive. This is because some journals in physics request authors to submit their manuscripts through arXiv. Authors have to submit their manuscripts to arXiv and notify the journal, then the journal takes the manuscripts from arXiv and puts it in the review and publication process, therefore all of the articles that are supposed to be submitted to some journals have their preprint versions

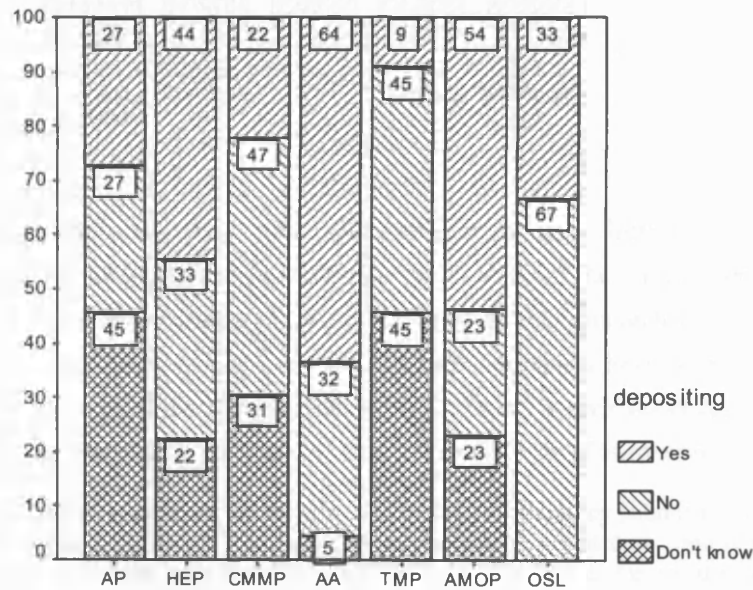
available on arXiv. This is something that a reader in HEP explained, when he was asked whether he deposited his papers in arXiv

Yes we submit it to the... well yeah, because actually if you want to submit it to Physics Letters or any journal it has to have gone to the archive first because they then upload it from the archive. So the publishers, actually the journal actually uploads it from the pre-print server as well. [AHE4]

The lowest depositing rates belonged to Theoretical Molecular Physics (9%) and Condensed Matter and Material Physics (22%). These variations among different research groups were also evident from the outcome of the interviews. An interviewee from the CMMP group said that preprints are not very popular in his group.

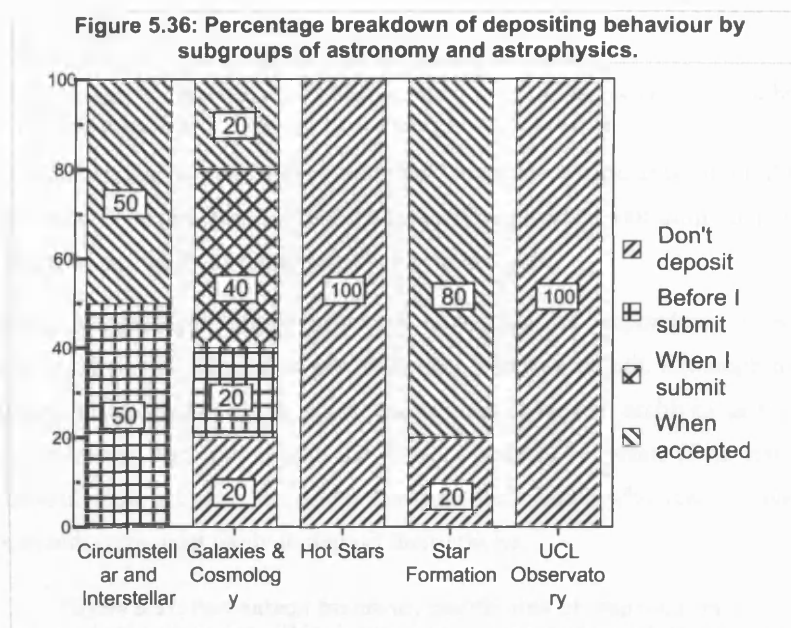
The preprint archives are not really... is not really an active operation in condensed matter physics. I'm thinking at the moment with a colleague of sending something to pre-print archive, largely for a tactical reason, but I don't look at the pre-print archives. [ACM1]

**Figure 5.35: Percentage frequency distribution of respondents by whether they deposit their articles by research groups.**



To understand the interaction of users with e-print archives even better, 22 respondents in the Astronomy and Astrophysics group were broken into their main six subgroups. Figure 5.36 reveals that while in some subgroups of Astronomy and Astrophysics such as Observatory, or Hot Stars, respondents did not deposit their articles at all. In some others such as Star Formation they tended to deposit their articles after the acceptance of articles in the journals. Two-fifth of respondents in Galaxies and Cosmology deposited at the same time

when they submitted their articles to journals. Respondents in Circumstellar and Interstellar usually deposit their articles before they are submitted to journals.



These findings approved what the interviewees said. Interviews also showed differences among different sub-fields of astronomy in terms of their interaction with e-print archives. While people in Cosmology deposit their papers in e-print archives as soon as they submit to journals, people in Hot Stars deposit once the papers have been accepted. The reason interviewees stated for this early submission of papers to archives in cosmology is that their field is more competitive compared to some other subfields of astronomy.

The cosmology group they submit it to Astro-ph as soon as they submit to a journal and their reasoning which I think is right in their field, is that their field is just so competitive. If you wait until the paper is being published, the results will be superseded by ten other groups, so they have to, in fact I was told by some cosmologist that some of them actually submit it to Astro-ph even before submitting to a journal, so as soon as they put together the results, they just submit it because they need to get the results out in the community, so in their field I think it's a bit more dangerous say for a student, because a student may not know the different groups. Some download something from Astro-Ph and it could be wrong, because it hasn't been refereed at all, nor even submitted. In our field I actually don't know anyone in star formation who submits as soon as they, submit it so Astro-Ph as soon as they submit it to a journal, although I have been told that people are starting doing that so we may be where we are all headed, which means in a way that Astro-Ph is replacing journals, right, so it may well be in ten years' time we all do, but so far certainly it depends on the field that you're in. [AAA1]

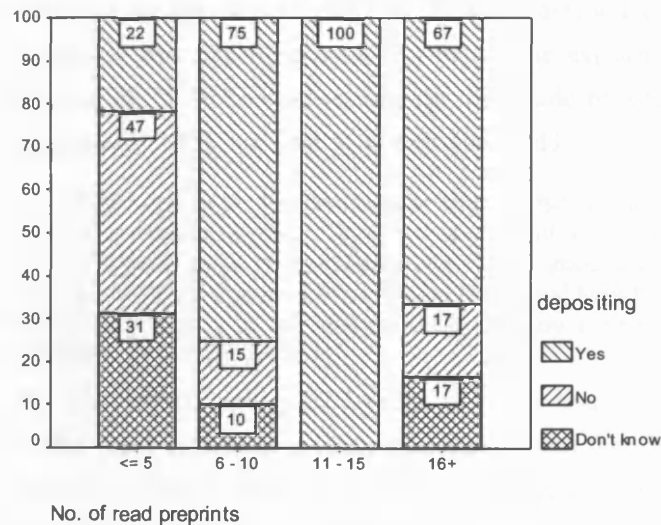
A professor also stated the same reason about early submission by cosmologists by saying that they seem to feel that they need to get stuff out immediately [AAA5]. Another professor [AAA4] considered a cultural element in this variation. This refers to the culture in a research community. He also maintained that the older people at his age would be different from younger people.

I think younger workers in my field [hot stars] won't think twice of just submit to Astro PH automatically. Older ones probably won't. [AAA4]

This difference is because older people tend to be more concerned about the validity of the data and the peer review process therefore they might tend to wait until the paper has been accepted and then put the final version on the e-print archive.

Figure 5.37 gives percentage frequency distribution of respondents by whether they deposited their articles by number of preprints they read in a month. Although there is not a steady increase in the likelihood of depositing articles in e-print archives as the number of articles read increases, those who read the fewest number of e-prints (5 or less) were least likely to deposit their articles, just about two-fifth did. Those who read in average 11-15 preprints a month were most likely to deposit their articles.

**Figure 5.37: Percentage frequency distribution of respondents by whether they deposit their articles by number of preprints they read in a month.**



## 5.10. Problems and difficulties

One of the objectives of any study in the field of information-seeking behaviour is to find out ways to improve information-seeking techniques and enhance information services. This is not possible unless we know about the barriers and difficulties users actually encounter or may

encounter in their use of services and during their information seeking activities. This section presents the findings of the study about problems and difficulties of users in their information seeking activities.

The data presented here are from interviews, information-event cards as well as the survey. All of the interviewees in this study were asked about the problems and difficulties they faced in their information seeking activities. Information-event cards also had a section that users could use to report any difficulties or unusual incidents they faced during their information seeking activities. Two of the questions in the questionnaire survey were pertinent to problems. A multiple choice question with five options asked respondents about their problems. Another question asked respondents about their level of agreement or disagreement with two statements on the access to journals and conducting subject searches. The data obtained through all of these three methods are organised here under different problems.

The interviews showed that about 60 percent of the interviewees were relatively happy with the level and ease of their access to scholarly information, particularly scholarly journals, and with the information services they are provided with. This is because the UCL library is subscribed to a considerable number of journals and provides access to most of the important scientific databases. One of the interviewees appreciatively stated that 'having access to the UCL electronic resources makes life very easy.' [ACM1]

The other reason for the general satisfaction of the participants in the study, especially for the older interviewees who experienced the era before the expansion of networked information systems and the Web, was the comparison they made between availability of information today and in early 1990s and before. As a professor said:

I think it's remarkably easy in the last decade or so to get information compared to when I started as a research student. Then it was difficult to get information because it was all in paper form. If your library wasn't very good, and if it didn't have conference proceedings or journal subscription, you would have to go to a big lab somewhere and try to get it. It was really difficult. But now it's so easy. There aren't any problems I would say. [AHE3]

However, the interviewees highlighted a range of issues that concerned them and caused them difficulties from time to time. They highlighted areas that they wanted to be improved. These issues were mostly related to access to information resources. Table 5.12 lists the difficulties and problems of the participants based on their comments in the interviews. The numbers represent the number (out of 56 interviewees) and percentage of the interviewees who mentioned the problem.



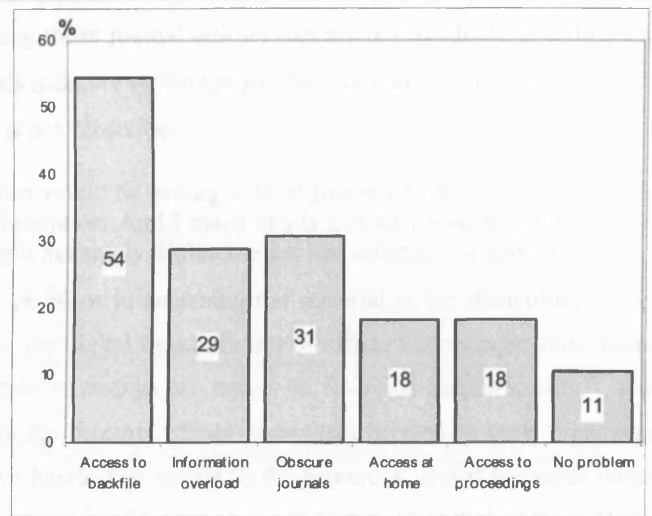
**Table 5.12: List of problems and difficulties of participants in their information seeking activities.**

- Problems in accessing information resources, 10 (18%)
- Difficulty accessing older articles, 6 (11%)
- Difficulty accessing obscure journals, 4 (7%)
- Difficulty accessing journals and other electronic resources out of office, 4 (7%)
- Difficulty accessing conference proceedings, 3 (5%)
- Difficulty dealing with too many publications and information overload, 5 (10%)
- Finding information, 12 (21%)
- Lack of technical information and difficulty finding it, 1 (1.7%)
- Difficulty finding minor pieces of information in scholarly information resources, 1 (1.7%)
- Inefficient subject and keyword searching in some databases, which results in too many hits, or does not result in relevant material, 8 (14%)
- Ambiguity in information surrogates including ambiguous titles and abstracts of articles, 2 (3.5%)
- Lack of time to keep up with the publications and developments, 17, (30%)
- Personal information management problems, 1 (1.7%)
- Loss of serendipity due to diminishing of print journals, 9 (16%)

Based on the issues raised by the interviewees (both staff and students), a question was included in the survey and respondents were asked if they had any problems with access to the backfile of journals, access to obscure journals, access to journals at home, access to proceedings, and information overload. The results of the questionnaire survey confirmed some of the findings of the interviews. Slightly more than half (54%) of the respondents experienced difficulty accessing older issues of scholarly journals (Figure 5.38). Lack of access to obscure journals that the university does not subscribe to was the second common problem with 31% of respondents experiencing it. About a tenth of respondents in the survey maintained that they did not have any problems with their information seeking activities. Besides the given options, the respondents were provided with a textbox to add any other difficulties they might have faced. One respondent mentioned the lack of access to commercial publishers' databases with restricted access as a problem, and another one saw his/her inability to find relevant information as a difficulty. A third of respondent complained

about being located far from the UCL library and therefore having difficulty accessing those materials that are available just in hard copies such as conference proceedings.

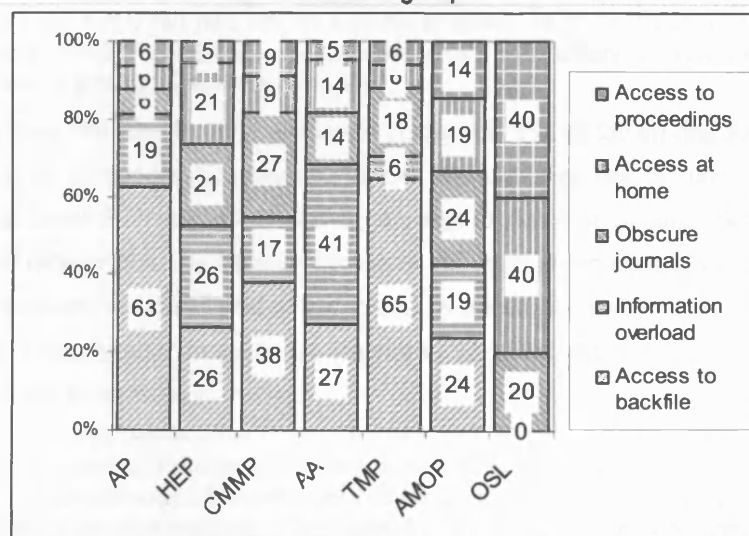
Figure 5.38: Percentage frequency distribution of problems.



\*Note that the respondents were able to choose multiple options and the cumulative percentage is more than 100 percent.

Different problems had different frequencies among different research groups. Figure 5.39 shows the relation between respondents' research groups and the commonality of the problems. Lack of access to the backfile of journals had the highest frequency among problems in the Theoretical Molecular Physics, Atmospheric Physics, and Condensed Matter and Material Physics research groups. In Astronomy and Astrophysics, information overload was more common than the other problems.

Figure 5.39: Frequency distribution of problems by respondents' research group.



### 5.10.1. Access to information resources

Access to information was something that ten (18%) of the interviewees mentioned as a problem. The access problem has a few different aspects. The main aspect was having difficulty accessing those journal articles that are not available in online format, or accessing articles in journals (obscure or foreign journals such as Japanese or Russian journals) to which the UCL library is not subscribed.

The problem would be getting hold of journals to which your institution does not have a subscription. And I mean this is getting worse and now certain publishers are making it extremely difficult to get free information. [ACM1]

The other problem in accessing the material is the difficulties some users face when they try to access the digital resources out of office, for example from home. Access to most of digital resources nowadays are based on IP-based authentication<sup>18</sup>, therefore users who attempt to access the resources from computers located in their organisations (in this case UCL) would have hassle-free access to the resources. But if the same users try to access the resources at home they would need to log in using either their Athens IDs or their institution IDs. This is considered by some users a bit of hassle.

Generally it appears that astronomy is richer in terms of digital information compared to physics. Almost all of the articles and journals in astronomy are electronically available and most of them have their backfiles digitised back to the first issue of the first volume. One interviewee [ATM1] particularly expressed his happiness with the availability of older issues of journals in astronomy. However, these old papers that are usually digitised through scanning sometimes cause problems as they tend to be large files and slow to download.

The only problem is old papers because they've been scanned in and they tend to be so huge that it can take you 45 minutes to download it. That's about the only problem. Realistically actually, no I mean there's... I would say we have very little problem in getting information. [AHE4]

This open and friendly information environment is one of the driving forces of a new research era in astronomy. Creation of Virtual Observatories has brought a new era in astronomical research. Virtual Observatories are repositories that contain observational and astronomical data collected by different telescopes and other observation devices. Their aim is to allow transparent and distributed access to data available worldwide. Now it is technically possible for anybody with access to the internet to use these data and reinterpret them and conduct research in astronomy.

---

<sup>18</sup> IP-based authentication is a method that institutions such as universities that use a fixed Internet Protocol (IP) address or range of IP addresses through which they connect to the internet use in order to give access to all of their users. This method is used as an alternative to using password and it is mainly used because of its convenience.

A lecturer in Astronomy [AAA1] complained about introvert people who want to keep the information for themselves. She believed this hinder the access to scholarly information and results in repetitive work as one researcher may end up doing something without knowing the fact that the same has already been done by someone else. This reluctance for disclosing scientific information can sometimes be related to the nature of the field and how tied the field is to the industrial and private section. The astronomers who were interviewed believed the information environment in astronomy is more open in this regard compared to information environment in physics, especially certain areas such as optics that have industrial applications. As one of the interviewees explained while in some areas such as designing optical systems, which has commercial application, scientists would not want to disclose their information, in astronomy it tends to be like 'do you want to chat, ok, I've got ten minutes, what do you want to chat about?' [SAM4]. Two other interviewees in astronomy also made similar remarks.

in terms of information services and access to the journals it seems to work fine. I think it helps that no one in astronomy is doing anything commercial. [SAA2]

In astronomy everybody is very happy. I suppose it's different. I know people working in chemistry and biology and there is a business behind them and they're paying for the research to be done and they want research because it's going to make their money whereas in the astronomy they're paying for the research because research should be done. And whether me and my supervisor doing it here or people doing it in Manchester or anywhere else in the world we all want to produce the results. The one thing that people don't like is people who work on the exact same thing because then somebody else might get the results out before you. [SAA1]

A few (5 or 19%) of participants in information-event card study reported their difficulties. The small number of problems reported in the cards reflect the good rate of the successful information seeking events. However, among the reported difficulties, access was the main issue. Three users wrote about the difficulties they faced accessing electronic journals using SFX<sup>19</sup>, accessing older articles and backfiles of journals. One user complained about lack of access to backfiles of journals.

all or most of the papers before 1990 are not yet in electronic version, this increase the searching time from 5 min to 15 min [SCM4]

### 5.10.2. Information overload

Information overload as defined by Bawden is:

that state in which available, and potentially useful, information is a hindrance rather than a help. It is associated with: loss of control over information, inability

---

<sup>19</sup> SFX is OpenURL link resolver. It is a tool for interconnecting library-controlled resources and services. SFX allows context-sensitive linking between Web resources in the scholarly information environment.

to use information effectively, and inefficient work and possible risk to health (2001:6).

Some of the difficulties that the interviewees mentioned could be categories as information overload based on this definition. Having to deal with too much information and being overwhelmed by too many scholarly papers and publications was an issue that was raised by eight of the interviewees. In the questionnaire survey, 29% of the respondents also said they had difficulties with information overload (Figure 5.35). About ten percent of the interviewees seemed to find it hard to keep up with so many papers that are published on a daily basis in their field. One of them for example stated:

There seems to be so many papers and it's really hard to keep up with. And I'm a sort of broader in my subject so I could read so much; I have to decide 'look I can't read that, chuck it out'. That is the difficulty I am facing. I have been told it didn't use to be like that in the old days and old days mean until beginning of the nineties. And I know from the older colleagues they're a sort of rely on their students or postdocs telling them what's going on right now. I mean they have a sort of global view but they don't look at papers. But then there are other older colleagues who are always up-to-date and they always know new things. But this is a sort of difficulty I am facing, you know keeping up-to-date in the information age. Because publishing is so important these days there are a lot of papers that each is a minor step from a big ground breaking paper. [AAA3]

The sense of being overwhelmed by masses of papers and information seemed to be stronger among people with fewer years of experience, people who are at early stages of their research careers including research students and researchers. A researcher [ATM1] pointed out 'it can be a pain reading through Astro-ph every day because there are 30, 40 pages on it everyday'. However, there are ways that participants seemed to have adopted for overcoming this problem. For example one [ATM1] resorted to newsletters such as *COOLNEWS*, a monthly electronic newsletter whose purpose is to rapidly disseminate new research results dealing with cool stars and the sun. Another researcher's [AHE4] strategy for avoiding this problem was to rely on conferences and word of mouth. A reader [ACM4] stated that he did not use alert services, or did not browse table of contents of every single journal issue because there are too many of them. He preferred to take the view that serendipity would lead him to the interesting and relevant papers. Another interviewee hopelessly believed that 'there is no cure for that. You just have to spend time going through it.' [ACM10]

People with higher status such as senior researcher, reader and professor have years of experience and tend to have a better ability for choosing the material they would want to read. As a professor [ACM3] stated one way of dealing with this mass of information is 'to be more choosy' and this is something that 'old man can do it, but a young man probably can't quickly.' A lecturer [AAA1] stressed the same strategy of being selective by proposing two ways for dealing with this problem, which are reading very specialised, and being organised:

I think that is an unconscious filtering that one does in that you tend to read, you tend to specialise a lot, so you tend to read mainly papers, really on the sub-sub-sub field of the field that you are doing. That's one way, which is not necessarily right because unless you have new ideas you should expand. The second thing that I will say is by being very organised in that, for example our group we have weekly meetings where in a routine, two of us review papers they found [AAA1]

This necessity for being selective has another aspect. There was a belief among the participants that in today's information environment with massive numbers of papers being available on every subject the issue is not to find something but to make sure the best thing is found. A student [SAM4] articulated this notion while explaining the right strategy for dealing with masses of information. He explained that the current situation with mass availability of electronic information is like not seeing the wood for the trees. Therefore the right strategy is to develop a sense of judgment to decide what is best to study and be very selective

I think it's the case of developing the judgment, yes I think it's developing the ability to acknowledge what is good and bad information as opposed to just finding something in the old days would've been 'yes I found something, I can use it' whereas nowadays is 'am I using the best thing?, am I using the thing that is most appropriate? How do I make that judgement?' So that can be a major issue. [SAM4]

However, one has to appreciate that developing this judgment ability and the skill to find the best sources of information is time-consuming and needs experience and practice. One really needs to 'root around for what is the best.' [SAA3]. Generally some of the techniques that the interviewees mentioned they used for dealing with the information overload were the same as the solutions Bawden (2001) suggested including quality filters, personal information management and information literacy skills which can all be called personal effectiveness.

Three of the interviewees believed that this large number of articles is the result of over publishing trend. And over-publishing in its turn is the result of the existing pressure on scientists for publishing, as their promotion and research funding rely on their publication status. One interviewee stated that 'no publications no money, no money no research. It's crazy.' [ACM10] This is conceived by many scientists [e.g. ACM5] as a kind of existing publication policy generally in science. Basically scientists are forced to publish in good journals and very quickly. For example conference proceedings were maybe a reasonable source of information some twenty years ago, but now people just do not waste any result on that kind of publication. This trend has certain disadvantages of course. For instance scientists have to publish very fast and they have to maybe split up results. So while twenty years ago it was of course easier because people maybe published one paper and took their time till the project was finished and then presented the whole research in an appropriate journal, nowadays one project may result in tens of papers, each containing a bit of the findings.

... there is so many papers and a little bit of information in each paper, to get to the gist of that you have to look at so many papers these days. Twenty years ago maybe 20 papers would have been in one paper. But unfortunately that how the system works at the moment. You have to publish and get citations. [AAA2]

This over-publishing situation of course is not the same in all subfields of physics. As a experimentalist reader in high energy physics (HEP) remarked, the problem is more sever in theoretical area; and theoretician physicists tend to publish more papers.

...I think there are probably too many theoretical papers, but that's the nature of these people who either to progress in their careers or whatever need to publish quite frequently. The publication process can be longer in experimental HEP right, because we have to build the experiment, take the data, analyse the data, get it through our internal procedures. You know I've worked on one analysis which has taken 3 years and has produced one paper. You know for theoretical people also... that can happen theoretically, but often they can also do things where they do one... two or three things and they can produce five or six papers quite quickly. But our experiment produces forty, fifty papers a year and that has 800 names on it. [AHE4]

This opinion was also reflected in two international surveys of authors conducted by CIBER in 2005 and 2006 (Nicholas et al., 2005c, Nicholas, Jamali and Rowlands, 2006)

### 5.10.3. Technical information and minor pieces of information

A professor raised an important issue which is the difficulty of finding small pieces of information in the scholarly information environment. Finding small items of information that are important but are not the subject of articles can be problematic. In conventional classic databases that the information retrieval is based mainly on the title, abstract and keyword search, it is hard to find this kind of information. But with new giant search systems such as Google Scholar that make it possible to perform full-text searches, it is more likely to find this kind of information, even though they might not be very efficient. The interviewee also suggested a solution to this problem based on his personal experience.

...I can give an example and a possible solution. One problem has always been the properties of materials. Simple properties like what's the elastic consent of a particular metal or a dialectic consent of some particular insulator. Now that's very scattered in the literature and normally because it is never the subject of the paper in which it was measured. ... Now what I did about it having got frustrated, I built up a data base of those myself just in a large book. And now in fact that is available electronically. The small company run by a friend of mine, Oxford Materials, has my database on there on the web. Now I personally will go to my shelf, but I mean other people can go to the web. [ACM3]

Finding technical information is another difficulty. Some technical information is not presented in journal articles, they are sometimes documented in the internal papers of the research groups for their own use or they are published on a Website. This sometimes causes difficulty in some areas of physics such as optical systems, where commercial companies

involved. A research student from optical laboratory [AOS1] explained his difficulties finding this kind of technical information.

In observational astronomy this problem manifests itself as lack of documentation of observational techniques. While people develop their own techniques for working with the observational devices, there seems to be no proper mechanism (in this case at least at UCL) for documentation of the methods and techniques and facilitating the use of devices by novice researchers and students.

The hardest thing I found is observation techniques. Because having gone to the telescopes, we need to know how to turn those essentials into a number. Somebody has to sit and write the definitive guide because everyone seems to have their own convention, their own units and their methods. It's a complete nightmare trying to work out how to do this. It seems we are not very good at writing down anything practical. [SAA2]

#### 5.10.4. Inefficient searches

As the number of scientists and therefore publications grow year by year, it gets more difficult for scientists to memorise and remember the names of even the important researchers from different nationalities in their fields. The other side-effect of the increasing number of authors is the increasing number of homonyms (authors with similar names) that makes it hard to identify papers by specific authors and differentiate them from works by other authors with similar names in the same field. This appears to become more problematic as fast-developing countries particularly China, has dramatically increased the number of their scientists and their international contribution in science. China has become a major player in the production of scientific papers. Its contribution to world science shows exponential growth, which is unique in the world (Leydesdorff and Zhou, 2005; Zhou and Leydesdorff, 2006). A professor in astronomy found it difficult to search for authors in ADS and suggested to create and use a system similar to the Digital Object Identifier<sup>20</sup> (DOI) for authors.

You know in Chinese there are seven surnames shared between 800 million people, so Lius and Fongs and you know, it just is so many. So even with initials going in like X. Liu, you still get so many. So that is a big problem, and that is going to be a bigger problem because more and more Chinese are growing up, becoming scientists so that means I got a collaborator called Xiao Liu. He is a scientist in Beijing. He is a professor, so I put an X, let me try X. Liu. It turns out that they are not all by him, so when you can't always tell which one is which you know, you have got to go through them by hand. That is a problem, so they are going to need, you know when you go into a paper they have got DOI identified, well I think

---

<sup>20</sup> It is a standard system for identifying content objects in the digital environment and associating it with related data. It is developed by the International DOI Foundation on behalf of the publishing industry. A typical use of a DOI is to give a scientific paper or article a unique identifying number that can be used by anyone to locate details of the paper, and possibly an electronic copy.



authors are going to need DOIs at some point and you basically put in there DOI because it is getting very difficult for some of these authors. ...so Chinese are problem. I mean Smith in English and Jones. The Welsh share the problem as well. Because there are not many surnames in the Welsh, Mooney Moons, Davies, Jones, they are very, very common so you get five, five or six different. R. Davies is in Britain there are four R. Davies astronomers. [AAA5]

The other problem with searching for authors is that a lot of people rely on searching for papers by key authors in the field, and as a few of the interviewees pointed out [e.g. AAA2 & AAA8] this includes the danger of missing out good papers by less known authors.

Information retrieval is an important and large area of research with many researchers from computer science and information science researching it from different perspectives. As advanced as they have become, the information retrieval systems are still far from the ideal and yet to be improved to satisfy users' needs (Allan, Carterette and Lewis, 2005). Every database and information retrieval system has its advantages and pitfalls. Inefficient information retrieval, from the perspective of users, was one of the problems that participants in this study mentioned. One interviewee [AAA6] was not sure that her subject searches using keywords result in all of the relevant and recent papers. She phrased her problem as 'being worried about the paper that you've missed'. This inefficient keyword searches have made scientists reluctant to use search email alerts. Databases such as Web of Knowledge and also arXiv have a feature that enables users to receive regular email alerts which include search results on the key words they specify. Two of the interviewees mentioned that they have tried such services for keeping up-to-date but later abandoned them because they realised the service was inefficient.

Another interviewee [AAM1] also raised the issue of inefficient information retrieval mechanism by complaining about the search options of databases such as ADS and Institute of Physics (IoP) as well as arXiv.

Indexing languages and the problems they may cause in subject and keyword searches was another aspect of information retrieval that was mentioned by a researcher [ACM5]. He found searching for keyword inefficient and opted to seek the literature by following references rather than by conducting subject or keyword searches. According to him, the problem with the keywords is that somebody somewhere defines the keywords and it is likely that they are not a good choice. For example people may not be very experts in the field. The other aspect of the problem with keywords is use of PACS<sup>21</sup> codes, which authors are

---

<sup>21</sup> The Physics and Astronomy Classification Scheme (PACS), which is prepared mainly by the American Institute of Physics (AIP), is a hierarchical subject classification scheme for physics and astronomy. It provides an essential tool for classification and efficient retrieval of literature in physics and astronomy.

supposed to use when they submit papers. Because fields and subfields and the natural vocabulary used by the researchers develop faster than the official classifications such as PACS do, then finding some information using those keywords will be difficult.

The problem that concerned this researcher is an old and still ongoing debate (see Rowley, 1994) in information retrieval which deals with advantages and disadvantages of use of controlled vocabularies and thesauri against the use of natural language (uncontrolled vocabularies). While natural language tends to be more up-to-date and less efficient in terms of precision, controlled vocabularies are slower in terms of being up-to-date and better in terms of precision. Another researcher [AAA2] also highlighted this problem with keyword search and using PACS. He thought that the use of more generic keywords will result in a lot of papers, and there are problems with using PACS. Apparently when authors want to assign PACS keywords to their papers, a lot of times they just copy paste them from the past papers. This makes the systems less efficient to search for more specific topics.

A helpful method of finding relevant information is to search for authors who are active in a specific field and track their publications in order to keep abreast with the development of a field and find key papers on that subject. However the problem with this method, as a participant mentioned, is to find out who the key researchers in an area are. This might be more difficult for newcomers in a field, people such as research students and novice researchers who have not been active in a field for long enough to know its research community. The mere reliance on author search also has the disadvantage of missing papers written by less known authors.

The hardest thing, I suppose, is simply to know who are the main players and this of course is a matter of detective work at the very outset, if it's a new field, detective work but once you know who they are it's incredibly easy to get hold of papers. [ACM4]

One interviewee [AHE5] was particularly unhappy with the search feature of two of the most important information resources in physics, which are Spires and arXiv. The problem appears to be caused by the way Spires and arXiv index the papers. In high energy physics every experiment has a name by which it is known among the HEP community. Each experiment also has a number, which is not necessarily memorised or known by the HEP community. The databases use the number for indexing the papers written about each experiment while users tend to use the experiments names to search for the relevant papers. This causes a problem as users who use the names are unable to find what they need. The other reason for this problem might have something to do with when papers are submitted and if Spires can specify correctly which experiment they correspond to.

As some of the interviewees were unhappy about the efficiency of searching databases for finding articles, a statement was included in the questionnaire survey and they were asked to express their level of agreement or disagreement. The statement was: 'I tend to avoid using subject keywords and phrases when searching databases to find articles because it brings up too many results.'

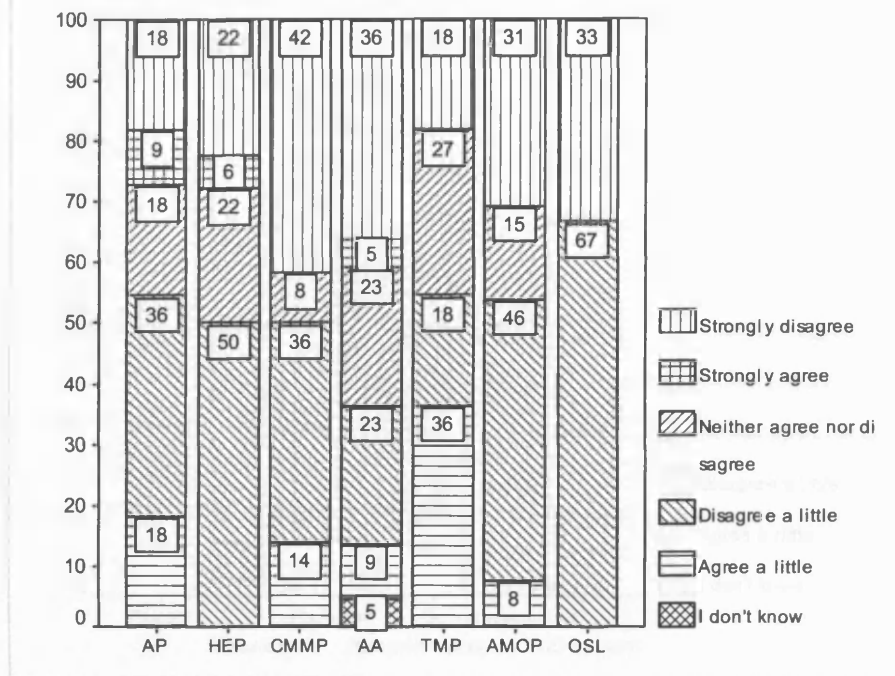
The majority of respondents (58%) were a little or strongly disagreed with the statements, however, 15% of respondents agreed a little or strongly that they would avoid conducting subject keywords searches. Table 5.13 relates.

**Table 5.13: Distribution of respondents by their level of agreement with the statement: 'I tend to avoid using subject keywords and phrases when searching databases to find articles because it brings up too many results'.**

	Avoiding subject searches	
	N	%
I don't know	1	.9
Strongly agree	3	2.6
Agree a little	14	12.3
Neither agree nor disagree	19	16.7
Disagree a little	41	36.0
Strongly disagree	36	31.6
Total	114	100

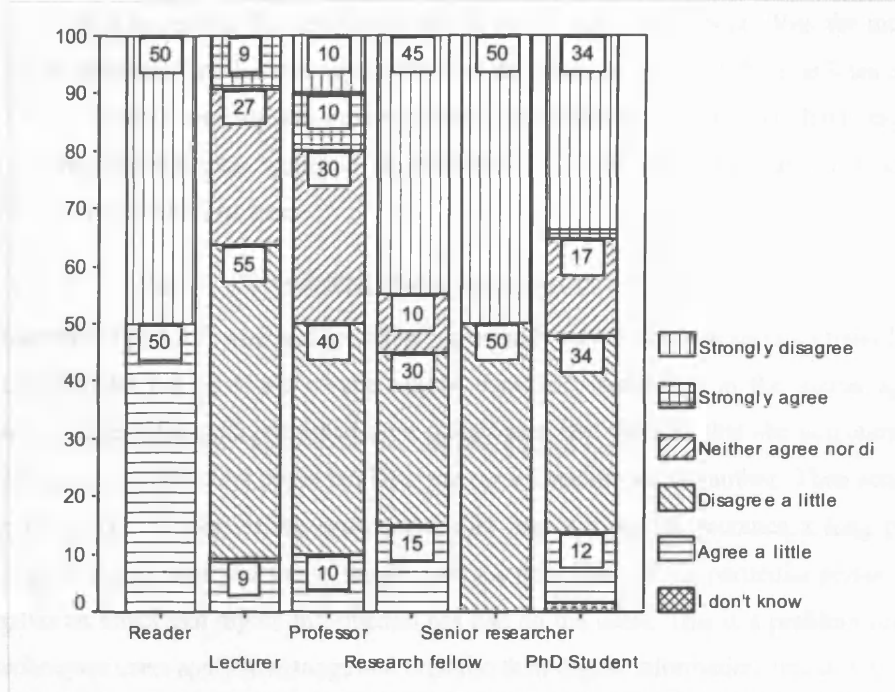
The respondents from different research groups were not very different from each other in terms of their agreement or disagreement with the statement on avoiding search for subject keywords. As Figure 5.40 shows, respondents in Theoretical Molecular Physics had the highest rate of agreement with the statement. Slightly more than 36% of them were agreed a little with this statement. Nine percent of respondents in Atmospheric Physics strongly agreed with the statement.

Figure 5.40: Percentage breakdown of responses to the statement on subject keyword searching by respondents' research group.



The highest agreement with the statement on subject keyword searching was among professors of whom a fifth agreed a little or strongly with the statement. The highest rate of disagreement belonged to senior researchers with all of them disagreed strongly or a little, and then research fellows with two-third of them disagreed strongly or a little (the figures for readers should be ignored as there were just two readers among respondents). See Figure 5.41.

Figure 5.41: Percentage breakdown of responses to the statement on subject keyword searching by respondents' status.



### 5.10.5. Ambiguity of the abstracts

Ambiguity in the publication and presentation of the research findings is another problem that one of the participants highlighted. It is self-evident that users do not read the full-text of every single article in order to decide whether it is relevant or not. Although due to the ease of full-text accessibility of scholarly articles more users now scan through the full-text of the articles to perform a relevance judgement, abstracts and titles are still the initial means for judging an article for its relevance to the topic searched (Nichols et al, 2007, abstract, JAL). If the title of an article or its abstract is not informative enough, users may be misled and this would result in a waste of time because users may end up reading parts of the article and after a while realising that the article is not pertinent to what they are actually looking for.

Also ambiguously. Sometimes you read, you can read an abstract where it appears to be something to do with what your interest is, and then when you read the paper there is nothing like you thought it was, so you have wasted an immense amount of time just trying to read it and then you find that is not actually going. [AAP1]

...and abstracts tend to be too vague. [SAA2]

### 5.10.6. Lack of time

Lack of time is something that seventeen interviewees complained about. With the increasing number of published articles that need to be read and pressure for publishing articles and also other duties such as teaching and supervising that an academic scientist may have, especially in the competitive British academic environment, it is not surprising that some scientists complain about the lack of time.

### 5.10.7. Personal information management

An interviewee [AAA7] who was generally happy with the information services remarked that her memory did not serve her as she wished especially nowadays in the digital age with massive number of papers published on a daily basis. She thought that she sometimes read something in one particular paper but later she cannot remember the author. Then sometimes when she needs to refer to the same paper and find it again, it becomes a long process. Although one may think that this is simply a personal problem of this particular person, in fact it implies an effect that digital information has had on the users. This is a problem related to the techniques users apply to manage and organise their digital information, and it is an area of research in information science. Of course this is a kind of problem that can be partly eliminated by utilising reference management software packages such as Endnote or Reference Manager. These software packages help users organise their references and their notes and they enable users to search in their notes. Improving personal information management skills also might help resolving such problems.

### 5.10.8. Browsing and serendipity

Mere reliance on online and electronic information has apparently resulted in the loss of serendipity, something that nine of the interviewees said they missed. Nowadays, few academics, especially in science are personally subscribed to scientific journals in print (Tenopir et al, 2003). None of the interviewees and the participants in the questionnaire study in this survey had personal subscription. The few who received print version of some journals did so through their society membership or because of their involvement in the publication of the journals as editors. The physical location of the physics and astronomy periodicals collection in the UCL library which is placed in the Science Library on the other side of the campus and relatively far from the Department makes users reluctant to go there for creative loafing and browsing. Four of the interviewees [e.g. AAA5, AHE5] complained that use of the library and browsing the hard copy of journals is impractical because the library is on the far side of the College and it is too far. Another interviewee [ACM1] said that he did not find

browsing on the screen very useful and maintained that 'browsing is something which is best done in a relaxed environment'. They consider it something time-consuming and therefore not worthwhile. As a result people in the Department do not have easy access to print journals and miss the advantage of browsing print journals. Other studies also showed the serendipitous discovery of information has diminished as a result of shift to electronic journals (Hallmark, 2004; Kurata et al., 2007). The following quotations show the interviewees' opinion about print journals.

It is worth mentioning that the purpose of browsing here is not keeping up-to-date, something that scientists used to do before 1990s. Nowadays, especially in physics where e-print archives are very advanced and popular, journals are considered 'out of date by definition', as a professor put it [AHE3]. According to another interviewee [ACM3] there are two reasons for going to a journal. One is looking for something specific and in this case the web is quite a good method. The other reason for picking up a journal is to be surprised. The same interviewee stated that the reason many like him read journals such as *Nature*, *Science*, *Physical Review Letters* is:

because of the article which are not in my field, they're ones which have surprises in for me. And to get those surprises which could really be very important for your research you are looking for things that you don't know about and the web is... well sure, you can let the web choose randomly, but I mean it will choose something very strange. It won't necessarily choose randomly in the areas you want. [ACM3]

The following two quotations are also related to the issue of serendipity and the desire for browsing.

...it is certainly true that I used to use, go to the library and look at the paper journals to get a broader perspective on what is going on, even within my own area. The only time I do that now is when I go to the Houseman<sup>22</sup> room and look at *Nature* or *New Scientist*. It is just much easier if you get a journal you know, you have got a how many page journal lumping on your desk once a month, you have a cup of coffee and you just flick through it, look at the titles and personally I used to find that much easier. [AAA3]

...I haven't seen a journal in print, a printed copy of a journal, for a long time and in the summer I was visiting somewhere and actually I saw some printed journals and I was amazed and I thought how nice it was to actually pick up the journals and scan the contents page. [AAA6]

A senior researcher used the analogy of analogue and digital radios to compare the older print-based information environment with the current digital information environment in

---

<sup>22</sup> Houseman is a common room in the main building of the UCL. It is used by members of staff as a relaxation and common room. Some general scientific journals and newspapers are provided there for reading. Its location is very close the Department of Physics and Astronomy.

which the serendipity element is somewhat missing. With the old fashion analogue radios, one gets exposed to things that one does not expect while tuning. One may find something accidentally. Whereas with new digital radios one presses the button and he or she already gets the programme. But if one wants to scan, it is not very practical with the digital radios. He maintained that the same is true about digital information environment, for example electronic and print books.

...Like if I give you a book you look and suddenly you look at something and you say oh this is interesting. But if this is on the screen, the same book of course can be on screen, then if you look for specific subject then it's fine; if you look for magnetism and you get magnetism. You could find magnetism in some other field. How do you find it if you don't know exactly what you are looking for. Therefore in your field it is very important to discover a way of connecting between the book and the screen. I don't know. ... I mean you put supernova and you get supernova but this is not the way for knowledge. It is harming knowledge. I am sure it is good for quick access, you press the button and you get the paper; but all the things that you cannot find, all the people that you don't know. If you don't know it then how do you find it. [AAA8]

He linked the lack of serendipity to the over-publishing which was discussed a few paragraphs above. Although the digital information systems seem to be good for searching and finding particular pieces of information, the browsing features are yet to be improved.

However, the only journal that is still relatively easily available in print and quite a few of the interviewees mentioned that they browse it on a weekly basis was *Nature*. *Nature* is also available in the Houseman room, for example, where members of staff can sit on a couch relax and browse it over a drink or cup of coffee.

Overall, the participants raised a few issues that concerned them. Some of these difficulties could be attributed to the information systems and services and some to the general scholarly communication system, and some others to the personal skills and issues such as information literacy. The highlighted difficulties are presented here as follows.

### **5.11. Summary**

This chapter presented the results of the study. The findings that have been obtained using three different methods including interviews, questionnaire and information-even cards were presented thematically. The chapter described the research process in physics and astronomy, the methods used by physicists and astronomers for seeking literature and identifying journal articles, methods used for keeping up-to-date, problem-based information-seeking behaviour, the issues involved in accessing information, changes in their information-seeking behaviour over time, reading and publishing behaviour, and finally problems and difficulties that participants faced in their information seeking activities.



Overall, the results revealed some variations and similarities in the information-seeking behaviour of participants from the different subfields of physics and astronomy. The interpretation of the results and theoretical discussion will be presented in the next chapter, which is the Discussions and Conclusions. The next chapter will include discussions about the findings, intradisciplinary comparison of the outcome of the study and it will revisit the research questions of the study.

## CHAPTER 6 - DISCUSSIONS AND CONCLUSIONS

### 6.1. Introduction

In the course of meeting its aims and to answer the research questions, this study examined the information-seeking behaviour and scholarly communication of physicists and astronomers. The overall purpose was to gain an understanding of how physicists and astronomers seek information and communicate; and what similarities and dissimilarities exist among physicists and astronomers based on their subfields, academic status and stages of research. The population of the research included PhD students and faculty members in the Department of Physics and Astronomy at the University College London. In this chapter the main findings and conclusions of the current study regarding the objectives and questions of the study are presented and discussed. The chapter also discusses the opportunities that arise from this study for further research in the related areas.

Each of the sections presented below (sections 6.2 to 6.5) refers back to one of the research questions.

### 6.2. Information-seeking behaviour of physicists and astronomers

The thesis studied the information-seeking behaviour and certain aspects of scholarly communication of physicists and astronomers with regard to three different factors including research subfields, academic status and stages of research. Below the findings of the research regarding each of these factors are discussed.

#### 6.2.1. By subfields

The Department of Physics and Astronomy at UCL has seven major research groups and research areas (at the time when this research was conducted) that were investigated through the study. Similarities and dissimilarities emerged among them with regard to the different aspects of their information-seeking behaviour and the areas of scholarly communication, specifically the publishing behaviour that was investigated in this thesis.

### **6.2.1.1. Astronomy and Astrophysics (AA)**

Astronomy and astrophysics deal with the study of celestial objects (such as stars, and planets) and phenomena that originate outside the Earth's atmosphere. AA is a self-contained area (compared to the other subfields) in that it relies mainly on its own concentrated literature.

People in AA have a high reliance on the e-print archive (arXiv). This is because arXiv is a valuable source of information for this field in that it is very comprehensive. It includes the preprints of most of AA articles. AA together with HEP are the two areas that were formed and developed in arXiv before the other subfields were added to the arXiv and also they are the heaviest users of arXiv among the subfields of physics and astronomy. However, as discussed in section 5.9.3 the use of arXiv varies among subfields of AA.

Most of the articles read by people in AA are identified through recommendations by colleagues, searches in databases and email alerts. We found that recommendations by colleagues play an important role in AA. Besides the general importance of human information resources in physics and astronomy as it was shown in the thesis (see section 5.4) it was found out that AA group at UCL had a very active information environment in terms of communication that facilitates interpersonal and informal scientific communications.

The study also showed that people in AA relied on browsing e-print archives, e-print email alerts and word of mouth in order to keep up-to-date. As argued, arXiv plays a central role for people in AA as a source of information and it is even more important as a means for keeping up-to-date. This is because of the preprints deposited in arXiv. However, as the study showed browsing arXiv was more popular than receiving email alerts from it. The study showed that several physicists and astronomers considered email alert services inefficient. The main reason for this was lack of specialised alerting services in that users normally have to choose broader subjects and therefore they receive too many items. Improving email alert services by implementing better filters for selecting items to be emailed to users and making them more specialised and narrower in terms of the subjects areas users can choose is something that information system developers need to look into.

AA was also the only group that made use of newsletters for keeping up-to-date. Newsletters proved to be popular among astronomers.

### **6.2.1.2. Atomic, Molecular, Optical and Positron Physics (AMOP)**

The AMOP physicists study the collisions of atoms and molecules with electrons and positrons, their interactions and manipulation using light, quantum chaos and the properties of ultra-cold condensed matter. In terms of interdisciplinarity of the field and the degree of

scatter of the literature of the field, this group was in the middle range compared to the other groups investigated in this thesis.

People in AMOP relied a great deal on using journal websites for finding articles they read. Searching Google was also the second most used method by AMOP researchers through which articles were identified. Browsing e-journals and browsing e-print archives were the two most used methods for keeping up-to-date with the developments in their field. Noteworthy, they read fewer preprints compared to researchers in the other research groups and relied more on journal articles.

### **6.2.1.3. Atmospheric Physics (AP)**

Atmospheric physics, as a subfield of physics, is the application of physics to the study of the atmosphere. Atmospheric physics attempts to model Earth's atmosphere and the atmospheres of the other planets using fluid flow equations, chemical models, radiation balancing, and energy transfer processes in the atmosphere.

Among the seven areas investigated in this study, this area of physics appeared to be in the middle range in terms of interdisciplinarity and scatter of literature. The findings of the study showed that researchers in AP relied on colleagues and reference tracking for finding articles they read. However, word of mouth, browsing electronic journals and e-print archives were the main sources of information in order of importance for keeping up-to-date. Additionally the AP researchers read more published journal articles than they read preprints.

Although the researchers in this field considered their literature to be somewhat interdisciplinary in that they make use of other disciplines' information sources, they did not rely on searching for finding articles as much as for example researchers in AMOP (another relatively interdisciplinary area in physics) did. This raises a question about the argument that the more reliance on the literature of several fields leads to more use of searching in reference databases (Vakkari and Talja, 2005). Vakkari and Talja (2005) considered the cross-disciplinary use of the literature as a measure for the scatter of literature, while the present research argues that these are two different factors (see section 6.6.).

### **6.2.1.4. Condensed Matter and Materials Physics (CMMP)**

CMMP is a vast subfield of physics with many smaller research areas covering a wide range of research from quantum information processing to biophysics and from superconductivity to radiation damage in materials. As the range of research topics covered by CMMP shows and as the results of this study reveal CMMP is the most interdisciplinary subfield among the

subfields of physics and astronomy that were studied in this thesis. The degree of scatter of the literature in CMMP is also the highest among all seven subfields that were studied.

The researchers in CMMP mainly used recommendations by their colleagues, conducting searches in general databases and tracking references for identifying articles they read. In addition they tended to browse electronic journals, rely on word of mouth, search databases and receive journals table of contents email alerts in order to keep up-to-date. E-print archive was not as popular in CMMP as it is in other subfields, hence the reliance on journals and methods such as ToC email alerts and browsing electronic journals.

#### **6.2.1.5. High Energy Physics (HEP)**

High energy physics, also called particle physics, is a very (the most) self-contained subfield in physics (Similar to AA in this regard) that concerns itself with the elementary constituents of matter and radiation, and the interactions between them. HEP physicists perceived the literature of their field to be a highly concentrated literature corpus and the use of literature of other disciplines was not considerable among its researchers.

Researchers in HEP relied mostly on searches in subject databases for identifying articles they read. HEP section of arXiv and Spires were the two specialised databases that are used by HEP physicists. The second most used method was searching in Google. The fact that Google was the second used means by which articles were found in the field of HEP might be because of high availability of open access material in HEP that makes everything searchable by general search engines such as Google. As explained in the results chapter (section 5.5.1) Google is not used intentionally for finding articles but it presents articles in the search results as users conduct searches. In order to keep up-to-date with the developments in HEP they mainly depended on browsing e-print archives, word of mouth and meetings. The number of preprints they read was larger than the number of journal articles they read.

#### **6.2.1.6. Optical Science Laboratory (OSL)**

Optical Science Laboratory is a research group in physics that studies the generation of electromagnetic radiation, the properties of that radiation, and the interaction of that radiation with matter, especially its manipulation and control. However, traditionally OSL at UCL is more concerned with observational astronomy and the technologies upon which it depends and therefore the research of this group is mainly of instrumentation type.

The study showed that OSL was in mid range in terms of interdisciplinarity while the researchers in the field perceived their literature to be concentrated. The field has a few specialised journals of its own. Researchers of this field relied on tracking references and

colleagues for identifying articles they read. The findings on this particular field supports Bates's (2002) hypothesis that scholars in high scatter fields use chaining and browsing as their primary search methods. Attending conferences and receiving table of contents email alerts were the two main methods on which they relied for keeping up-to-date with the developments in their fields.

#### **6.2.1.7. Theoretical Molecular Physics (TMP)**

Theoretical molecular physics is in fact a subfield within molecular physics, which merely concerns itself with theoretical research. TMP has an interdisciplinary nature with a literature body that compared to the other subfields of physics and astronomy was scattered. The researchers in TMP relied mainly on tracking references for finding articles they read. Browsing electronic journals and searching databases were the two main methods used for keeping up-to-date by the researchers in this group. Preprints were not popular in TMP and its researchers mainly relied on journal articles.

The following table (Table 6.1) summarises the characteristics of the seven subfields of physics and astronomy studied in this thesis. It presents the main methods used for keeping up-to-date and finding articles, the average number of articles and preprints read in a month, and the rank of the subfield in terms of the scatter of its literature and its use of the literature of other disciplines (interdisciplinarity).

**Table 6.1: Information-seeking characteristics of research groups.**

	Main method used for keeping up-to-date	Main method used for finding articles	Average No journal articles read in a month	Average No preprints read in a month	Interdisciplinarity of the field*	Scatter of literature*
<b>AA</b>	Browsing e-print archive	Search in subject database	5.9	8	★★	★★★★
<b>AMOP</b>	Browsing e-journals	Browsing and searching e-journal sites	9.8	5.9	★★★	★★
<b>AP</b>	Browsing e-print, e-journal, Toc alert, Word of mouth	Tracking references	7.3	1.6	★★★★★	★★★
<b>CMMP</b>	Browsing e-journals	Search in general databases	11.6	4.3	★★★★★★★	★★★★★★★
<b>HEP</b>	Browsing e-print archive	Search in subject databases	3.2	5.9	★	★★★★★
<b>OSL</b>	Conferences	Tracking references	8.3	8	★★★★	★
<b>TMP</b>	Browsing e-journals	Tracking references	7.5	2.1	★★★★★★★	★★★★★★★

\* The ranking for this part of the table has been done based on the answers of the survey respondents to the questions on their use of the literature of other disciplines and the scatter of the literature of their field. The larger the number of stars the more interdisciplinary and the more scatter the literature of the field.

Referring to the research question No.1 of the study (Are there any important differences between different subfields of physics and astronomy with regard to different aspects of information-seeking behaviour?) the findings presented here show that, although similarities exist, there are significant differences as well. As we can see from the table, the subfields of physics and astronomy are different in terms of their reliance on different methods used for keeping up-to-date as well as methods used for finding articles. In general, the study lends support to the findings of Brown (1999b) that showed high usage of citation tracking. It also confirms some of the findings by Nicholas, Huntington and Jamali (2006) that showed physicists compared to scientists in some other subjects were more likely to be browsers. However, as we can see different subfields are different and talking of physicists here might be over-generalising the data.

The findings of the present study raises a question about the past hypotheses proposed by Bates (2002) and Vakkari and Talja (2005) about the effect of the degree of the scatter of the literature of a field and its interdisciplinary nature plays in information-seeking behaviour.

Bates (2002) suggested that the use of chaining and browsing are the main search methods used by scholars in high scatter fields and directed keyword searching is the more effective method for finding relevant materials in low scatter fields. On the other hand, Vakkari and Talja (2005) argued otherwise and maintained that the increase in scatter increases only the importance of searching in reference databases. However, the subject categories used in Vakkari and Talja's research were very broad and included categories such as Natural Sciences that encompassed physics, chemistry, food industry and some other disciplines. This kind of generalisation is misleading as we can see in the present thesis that differences exist within single disciplines let alone among all scientific fields that are encircled under Natural Sciences. While in an area such as CMMP with a highly scattered literature searching in general databases is the main searching method, in another area with scattered literature such as OSL tracking references is the main mean for looking for articles. Although people in OSL use the literature of other fields they believe their literature body is reasonably concentrated. It should be noted that the use of the literature of other disciplines does not necessarily imply the scatter of literature. For example researchers in a subfield of physics might use the chemistry literature, but the chemical papers they use might be mainly published in one or two specific journals and this means that the literature they rely on is not scattered, but concentrated in a few journals. Therefore, the interdisciplinary use and the scatter of literature might be considered two separate issues, while in Vakkari and Talja's study the use of literature from several disciplines has been interpreted as the scatter of literature.

The results also show the difference within physics and astronomy in reading behaviour and the use of journal articles and preprints. For example people in AP, TMP and CMMP, which are subfields with high interdisciplinary usage of literature tended to read published journal articles and few preprints. While some subfields such as HEP and AA that are self-contained relied more on preprints and read fewer published journal articles. This raises questions about the development of e-print archives. It is not clear whether facilitating the cross-disciplinary use of literature in the e-print archives would encourage its usage by researchers who use the literature of several disciplines. This area merits further investigation.

### 6.2.2. By academic status

The second research question of the study was "are there any important differences in the information-seeking behaviour of physicists and astronomers according to academic status?" The results of the study presented in chapter five showed that differences exist among users with different status concerning their information-seeking behaviour. This is expected as the academics' and researchers' information seeking skills evolve during their careers. However, the nature of the differences needed to be investigated.

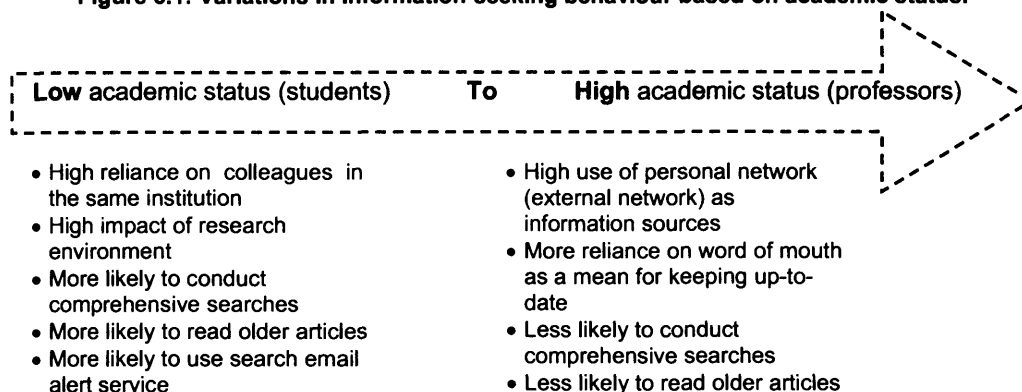


The study showed that PhD students relied highly on their senior colleagues as one of their important information resources. Supervisors and colleagues are the reference points for students in their information seeking activities. They are the ones who introduce PhD students to the subject area and help them develop their information seeking skills.

There was a resemblance between the students and the senior members of staff (e.g. professors) in that they also relied on their personal communications for their information seeking. The nature of this reliance is different. For PhD students the reliance is on the colleagues in the same room, neighbourhood or the same department while the reliance of senior colleagues on personal communication is based on their personal networks that they have developed during their careers. Thus physical proximity and the research and communication environment of the department has a much greater impact on the students' information-seeking behaviour than it has on the senior members of staff. The staff's network is not limited to the people in the same building or the department. This is because they have developed a wider and more sophisticated network during their longer careers. Therefore they take advantage of communication technologies to communicate with their peers who form their personal network around the world. The findings of the study highlight the importance of creating a department with a friendly research and communication environment as it helps students integrate in their subject community and conduct their research effectively. Students rely a great deal on their internal (i.e. departmental) network and their close colleagues as a valuable source of information and therefore physical distance and proximity matter for them. Elimination of the physical barriers for communication between peer PhD students and colleagues helps them enormously in their information seeking activities as well as in satisfying their information need.

The following diagram (Figure 6.1) compares how some of the key information-seeking characteristics change as academic status develops from low status (students) to higher status (professors). The lower the status the higher the likelihood of conducting comprehensive searches and reading older articles.

Figure 6.1: Variations in information-seeking behaviour based on academic status.



### 6.2.3. By research stages

Conducting a research project entails several information-seeking activities. The third research question of the study was “what are the techniques and methods applied to information seeking in the different stages of a research process in physics and astronomy” This study cast light on the information-seeking behaviour of researchers in physics and astronomy in different stages of a project and showed different techniques and approaches that physicists and astronomers used in the different stages of a research process.

The initial sparks and inspirations that lead to creation of a research project might occur during the process of acquiring information. Information obtained through different channels and by different information-seeking methods leads to the identification of a research problem or a potential research subject and ultimately a new research study is formed.

The findings of this study revealed that interpersonal communications (formal such as conferences and informal such as chats with colleagues) is of much significance in all subfields of physics and astronomy as a source of inspiration for new research projects. However, minor variations also emerged as to how new research projects are shaped in the different subfields of physics and astronomy. In theoretical physics as well as in astronomy reviews are a valuable source for identification of potential research questions while in experimental physics (especially HEP) theoretical papers function as a source of research questions. Theoretical papers suggest solutions for existing problems and experimentalists try to find ways to test the suggested solutions and therefore those theoretical papers are the basis of new research projects.

Once an initial inspiration is provided and the primary research problems on which a researcher or a research team would want to focus is clear, another stage of information seeking starts which is further investigation on the subject or the initial familiarisation.

Once more, differences existed in information-seeking at this stage. This stage of information-seeking involves high usage of research literature, mainly journal articles. However, researchers apply different methods to find relevant literature that they should read. The study showed that few people resort to conducting a comprehensive search in databases to identify all of the relevant literature. The initial approach for the majority of the researchers in physics and astronomy is to find some seed papers, or identify some key authority/experts on a given subject through communicating with colleagues and deploying their personal network. The degree to which personal contacts and colleagues are used at this stage varied among researchers in the different subfields of physics and astronomy. Furthermore, the study showed that organisational factors such as the physical proximity has impact on this. Once a few seed papers or key experts are identified then methods such as tracking references or searching for papers by the specific authors may be applied. However, a few people may resort to conducting subject searches as well. In larger groups this task of collecting and collating the background information is delegated to a few people and they inform the rest through internal documents and meetings.

During the course of a project, different methods are applied for keeping up-to-date. Generally one of the highlights of the findings of this study is the importance of interpersonal communication methods for keeping up-to-date among physicists and astronomers. Word of mouth and colleagues were among the favourite methods for keeping up-to-date, browsing e-print archives and e-journals were also used highly by physicist and astronomers for keeping up-to-date. However, when looking at specific subfields of physics and astronomy, again differences emerged. For example, in high energy physics the reliance on preprints was much higher than in the other groups such as the atmospheric physics. The subfields that are perceived to be more competitive and fast moving by its community tended to consider journal articles out of date for the purpose of keeping up-to-date while the subfields that were not as competitive relied on journals as they could afford to wait the time lapse between submission of articles and their publications in journals. There were also differences among people with different status or at different ages regarding the methods used for keeping up-to-date. The senior members of staff and the older ones relied more on their personal communications, colleagues and conferences for keeping up-to-date whereas PhD students and junior researchers were more reliant on the literature and methods such as email alerts for keeping abreast with the development of their fields.

In addition to the measures that physicists and astronomers take in order to keep their knowledge up-to-date, some problem-specific-information seeking instances are undertaken during the course of a research project. The study revealed some facts on the way physicists and astronomers go about problem-specific information seeking and their preferences for information resources. It was found that human sources and the Web are the most used and highly preferred resources for this kind of information-seeking. This is something that was also confirmed by a recent study (Savolainen, 2008).

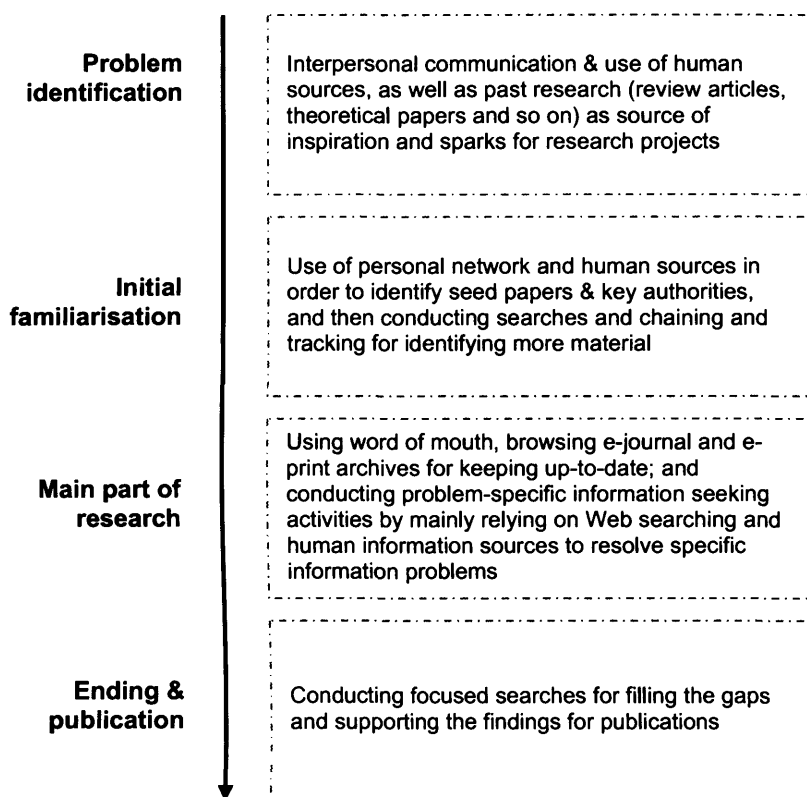
The study also answered the question about use of search engines by physicists and astronomers that had been raised by Nicholas et al. (2005b). Although search engines, and specifically Google play an important role in finding articles, it is not intentionally used for that purpose; however, locating articles is an outcome of the search. Noteworthy the current study found that physicists and astronomers are becoming more and more aware of this feature of Google searching and a change should be expected in their behaviour and attitude towards use of search engines for finding scholarly papers. As open access materials become popular and users learn to use search engines for locating scholarly articles, new information seeking paradigms might emerge and these emerging paradigms merit longitudinal investigation<sup>23</sup>.

Figure 6.2 illustrates the stages of a research project as described above.

---

<sup>23</sup> It is worth mentioning 'Google Generation' research project led by CIBER for the British Library and JISC here as it tries to discover: a) whether or not as a result of the digital transition and resources being created digitally, young people, the "Google generation", are searching for and researching content in new ways and if so, how this will shape the way they research and search in the future; b) whether or not new ways of searching and researching for content will prove to be any different from the way that existing researchers/scholars work. See: <http://www.publishing.ucl.ac.uk/behaviour.html>

Figure 6.2: Stages of information-seeking in a research project.



### 6.3. Reading behaviour

In answering the research question ‘what are the characteristics of reading behaviour of physicist and astronomers?’ the study revealed the characteristics of the reading behaviour of physicists and astronomers in terms of the quantity of reading (journal articles and preprints) as well as issues such as screen reading and the credibility of information. The main findings of the study with regard to the reading behaviour are presented below.

- In average, physicists and astronomers read eight journal articles and five preprints a month. In other words, they read 156 scholarly papers a year in average. The number of annual readings for physicists and astronomers varied in a number of studies. Tenopir and King’s (2002) study in early 2000s showed that scientists read about 130 articles per year across all work fields. Another survey by them in at Oak Ridge National Laboratory (ORNL), however, indicated that physicists read an

average of 204 articles per year, chemists read an average of 276 articles per year, and engineers read an average of 72 articles each year. Comparing Tenopir's figure with the outcome of the current study shows that the average number of articles read by physicists and astronomers at UCL is smaller than physicist at ORNL, but greater than the average of general scientists. This difference can be because of organisational differences. The population of this study were academics while ORNL is a national science laboratory. Reading behaviour also can change during the time because of the emergence of new technologies which more and less affect the quality, quantity and style of reading.

- There is a positive correlation between the amount of reading and the amount of writing, those who write more papers, read more papers too.
- The amount of reading is influenced by the type of activities academics conduct, those who spend more time teaching read fewer papers and those who spend more time doing research read more papers. This confirms the findings of some of the past studies (Tenopir and King, 1996, 2002) that reported a positive correlation between the amount of reading and professional performance and achievement, which are normally measured by the number of publications and grants.
- There is a negative relation between reliance on preprints and reliance on journal articles. Those who rely more on preprints tend to read fewer journal articles and those who tend to read journal articles read fewer preprints compared to the other groups. There are some variations in this regard among researchers in the different subfields of physics and astronomy. For example people in astronomy and astrophysics, and high energy physics read more preprints and far fewer journal articles, whilst people in theoretical molecular physics and atmospheric physics read more journal articles and rely less on preprints.
- The importance associated with having up-to-date information affects the amount of reading. Researchers who associate more importance with keeping up-to-date quite expectedly read more papers.
- The methods applied for finding articles affects the amount of reading. Physicists and astronomers who find articles through recommendations by colleagues read fewer papers compared to those who rely on mail alerting services, browsing and searching journals for finding articles.
- Recently published articles account for a large proportion of the readings, which implies that an important reason for reading is keeping up-to-date. This confirms the findings of Tenopir and King's (2002) study which showed that about 80% of

readings were articles less than one year old. This can be interpreted that an important reason for reading might be keeping up-to-date.

- Age and academic status have influence on the age of papers read. Those with higher academic status such as professors and those in older ages are less likely to read older articles compared to younger researchers. This might be due to the fact that senior researchers already know about the background, the history and the past literature of their subject and feel less need to read older articles compared to junior researchers. However, this is just a hypothesis and needs to be investigated further. The other reason might be the fact that younger researchers are more likely to use search facilities and as Nicholas et al. (2005a) argued and this study confirmed use of search facilities increases use of older articles.
- There are variations among the subfields of physics and astronomy regarding the age of papers read. People in astronomy and astrophysics tend to read more recent papers compared, for example, to researchers in atmospheric physics or atomic molecular physics. This also reflects the difference on the reliance of researchers on preprints and journal articles. Those who tend to read more recent papers are the ones who rely more on preprints and those who are more likely to read older articles are the ones who tend to rely more on the journals.
- The majority of the physicists and astronomers prefer to access journals electronically and at the same time avoid screen reading. The review of the past research by Jamali, Nicholas and Huntington (2005) also showed that scientists prefer mainly the PDF version and the reason tends to be that it is a print-friendly version and can be saved and archived electronically. However, this study showed that the preference for screen reading or reading a paper in print format depends on the nature of reading. Researchers print a paper if they want to read it thoroughly or to read while commuting. On the other hand, they screen read when they want to skim through the paper quickly, look for specific pieces of information and so on. It must be also mentioned that although the electronic version of journals are favourable by the majority for research purposes, print versions of journals are also in demand but for a different usage. For instance, physicists and astronomers prefer to have certain journals such as *Nature* in print format so they can browse in their convenience mainly for serendipitous discovery of the articles that might interest them.

Although the findings of this study discussed in section 5.8 improve our understanding of the reading behaviour of physicists and astronomers and highlights some of the

intradisciplinary differences, the whole picture of scientists' reading behaviour is currently somewhat blurred. Based on the findings of the present study and a series of past studies on reading behaviour of scientists, we know about the quantity of reading as well as the source of reading and the methods used for identifying the papers read. However, we know too little about the nature of reading phenomenon itself. Thus we need to obtain a deeper knowledge in this area so that the information services could be improved accordingly. Reiteration of prior studies such as the one by Bazerman (1985) could be a good starting point for further understanding of reading behaviour of scientists.

#### 6.4. Publishing behaviour

The fifth research question of the study was 'what are the characteristics of publishing behaviour of physicists and astronomers?' To address this question the study investigated some aspects of physicists and astronomers' publishing behaviour such as their interaction with journals and e-print archives as well as some parts of their information-seeking behaviour which relates to their publishing activities.

The study showed that during the course of a research, one of the two main literature searching tasks is conducted at the time of writing articles. Compared to the initial literature search at the beginning of a project, the search for the literature at the time of writing is less thorough and more focused. The reason is that researchers already know most of the literature relevant to their research by the time they reach to the point they want to publish their findings. This is the outcome of the search they conduct for the literature in the initial stage of the project as well as the measures they take to keep abreast with the relevant literature during the course of a project.

The other aspect of the publishing behaviour is the decision as to which journal an article should be submitted to. There have been a number of studies in this area in different disciplines and studies such as Gordon (1984) and Frank (1994) have investigated scientists' journal selection criteria. Gordon (1984) for example, reported that journals are mainly chosen on the basis of audience they reach rather than the reward they confer. On the other hand Frank (1994) found that journal prestige, most frequently published journal topics, and readership composition were the most important factors for initial manuscript submissions. However, this study showed that the criteria and how scientists come to their final decision about the journals they want to publish in is a complex issue and depends on many factors including the characteristics of that scientific field. The criteria applied for and the factors affecting journal selecting decision vary not only from one discipline to another, but also from one subfield to another in a discipline. As it was discussed in the results presented in section 5.9.2 number of journals available, page-charge policy of the journal, the composition of the



collaboration, the nationality of the authors, and involvement of one or more of the authors in a journal as editorial team are among the factors that could have bearing on the final decision.

The interaction of the authors with e-print archives is an issue that was investigated. Variations and differences were found among the subfields of physics and astronomy in terms of their interactions with e-print archives. Although more people in physics and astronomy deposit their papers in e-print archives, specifically arXiv, there are different patterns for doing this which are affected by the characteristics of the subfields. While in the fast-moving subfields of physics (such as star formation) the articles tend to be deposited after acceptance by the journals. In some others such as circumstellar and interstellar astronomy they are deposited before submission to journals and in cosmology they get deposited at the same time as they are submitted to journals. In some subfields such as hot stars people are likely not to deposit their articles at all, while the rate of depositing in some other areas such as high energy physics, and atomic, molecular and positron physics is high. Although the study revealed some factors such as the speed of developments in a field on these variations in interaction with e-print archives, this is another area that merits further investigation and could be the subject of a separate research itself. To the best of the researcher's knowledge no research has been done in this area to illustrate a clear map of the factors that influence on the scientists' patterns of interaction with e-print archives. The present study supports the finding of Wertman's (1999) study in that high energy physicists are among the top users of e-print archives. Wertman also found that condensed matter physicists were among top users of e-print archives while the present study showed otherwise. It must be noted that Wertman's study is already nine years old.

## 6.5. Problems in information seeking

The sixth research question of the study was 'what are the difficulties and problems that physicists and astronomers face in their information seeking activities?'. The study identified some of the difficulties that students and staff in the Department of Physics and Astronomy were facing in their information seeking activities. A list of them is presented in the fifth chapter (see section 5.10). As a good number of problems and difficulties were related to the access to journals. One of the problems is the difficulty in accessing journals outside campus. Use of Athens<sup>24</sup> allows researchers to access the electronic resources out of office. However, users seem to consider it a hassle to have to use their ID and Password for accessing the

---

<sup>24</sup> Athens is contracted by the Joint Information Systems Committee (JISC) to control access to web-based subscription services for UK further or higher education institutions. Athens authenticates and authorises users for access to online services such as Documents Online (see [www.athens.ac.uk](http://www.athens.ac.uk)).

resources. They want hassle-free and seamless access to the electronic resources. UCL has been promoting the use of CITRIX<sup>25</sup>. This application has helped access to all of the library services and resources and all of the computer application from outside campus. It simulates a campus-based computer on the user's machine and therefore all of the authentication for use of all resources will be IP based. This means while users are accessing the resources through CITRIX interface they do not need to use their Athens because it is as if their computer is located in campus and uses an authenticated UCL IP address.

A suggestion to help improve this situation is to try to provide a service where by a databases or search facility (such as Google Scholar<sup>26</sup>) that provides direct link to the full-text of the results (i.e. articles) is incorporated in library services. Once users access Google Scholar through their library, they will be automatically authenticated for accessing all of the resources that Google Scholar will lead them to.

Lack of access to backfiles of journals was another problem identified in this study. Although, publishers have been active in this area and many publishers have undertaken the digitisation of backfiles of their journal up to the first issue, this situation can still be improved.

Difficulty in accessing obscure journals or journals that are not for any reason subscribed by the libraries is another problem that needs to be dealt with. Improving the open access culture and depositing papers globally may help resolve this problem. If authors deposit the preprints or postprints of their articles in repositories, the access to the scholarly information will be improved globally. However, one issue that should be paid attention to is the different versions of the same article that can be found on different repositories as well as the published version in the journal. There should be mechanisms for the automatic identification of different versions of the same article. Although attention has already been paid to this issue (see for instance Frankel et al., 2000) and even research has been undertaken to investigate the issue<sup>27</sup>, little practical or technical effort has been made to resolve this issue.<sup>28</sup>

---

<sup>25</sup> Citrix MetaFrame Presentation Server is a technology that can allow remote users to connect to applications that are actually installed on a remote computer.

<sup>26</sup> <http://scholar.google.com>

<sup>27</sup> VERSIONS Project (Versions of Eprints – a user Requirements Study and Investigation Of the Need for Standards) is a research project that addresses the issues and uncertainties relating to versions of academic papers in digital repositories (see <http://www.lse.ac.uk/library/versions/>).

<sup>28</sup> Probably the only current mechanism that does somewhat something in this area is Google Scholar that offers a link called 'all x versions'. The link lists records of the same articles found on different locations; they might include the same article on different repositories, the publisher website, as well as vendors. Although it is a very poor mechanism for identifying preprints of published articles and matching them to the published version, it helps sometimes.

Low usage of the physical library and the complaints about unavailability of print versions of journals and loss of serendipity could be seen among the participants in the study. Unfortunately, the UCL library buildings are mainly old and unsuitable for a library. Although a lot of modernisation and refurbishment works have been done recently and continues to be done, still some problems exist, such as shortage of space for books and journals, shortage of suitable desks with power supply for use of laptops and so on. UCL has 15 different libraries. But the Science Library that is home to the physics and astronomy collection is on the opposite side of the campus to the location of the Department and this distance makes people in the department reluctant to use the physical library. The department has a small library which is located in the same building as the department. One suggestion that can be very helpful is to improve the use of this library. One way of doing this is to make new issues of journals in physics and astronomy available for a week or two in this library before they are permanently placed in the Science Library. This will partly resolve the complaints that some people in the department had about lack of access to print journals for browsing and loss of serendipity.

Some of the other problems such as inefficient or unsuccessful searching in databases occur for two reasons, one can be problems on the system side that should be improved by the databases and the other reason could be low information literacy. Unfortunately, there has not been research on the information literacy of people in the Department, neither has there been research on the information literacy of physicist and astronomers in other places. The only research on the information literacy of postgraduate students in physics (Brown, 1999b) is already too old to rely on.

A suggestion to improve the situation is that the subject librarian can develop a short tailor-made information literacy course for the people in the department - especially for the postgraduate students. The course can be used for introducing important databases and resources in physics and astronomy, for teaching how to conduct efficient searches in these resources, and inform users of which databases are the most suitable for specific subfields, also to introduce value-added features that most of publishers of journals and databases offer and how to make the most of these services. Teaching some personal information management and use of reference management software packages could also help resolve some of the other problems such as complaints about difficulties in searching and finding information.

## **6.6. Contribution of the study**

This study is the first of its kind to investigate information-seeking behaviour of scientists in a single discipline with intradisciplinary approach. The need for such an approach and this kind

of research has been mentioned by some other researchers in the past (Case, 2002; Kling and McKim, 2000; Fry and Talja, 2004), however, no research has been done. The present research highlighted the importance of looking into smaller scientific domains for better understanding of information-seeking as well as communication behaviour. The researcher believes that the present research will help foster comparative studies of small scientific fields.

Adopting an intradisciplinary approach and based on its findings, the present research raised questions about the studies that have focused on very broad subject areas such as Natural Sciences (see for instance Vakkari and Talja, 2005). The information presented on the intradisciplinary differences in information-seeking behaviour of scientists in this study is an indication of the danger of presenting misleading information if several disciplines are encompassed under one single domain such as natural sciences or physical science.

The study also showed that considering the scatter of literature by itself as a determinant of the information-seeking behaviour of scientists can be misleading especially if the subject areas considered are broad. The study argues that the level of scatter of the literature of a field cannot determine the searching behaviour of scientists by itself as Bates (2002) and Vakkari and Talja (2005) have argued. The data presented in chapter 5 showed in some fields with scattered literature users resort to chaining and browsing, as Bates (2002) argues and in some other fields with scattered literature users resort to searching as Vakkari and Talja (2005) argue. The study also showed that cross-disciplinary use is a different issue from scatter of literature. Here the issue that should be mentioned is the concept of perceived interdisciplinarity of the field and perceived level of scatter of literature of the field by its researchers. The way users perceive the literature of their field to be (more interdisciplinary, more scattered and so on) affects the method they apply for their information seeking.

The current study contributes to our knowledge of information-seeking behaviour of physicists and astronomers. There has not been a comprehensive study of the information-seeking behaviour of physicist and astronomers during the last 15 years, a period of significant changes in the realm of information technology and services. The study updates our understanding of the information behaviour of this group of scientist who are good examples of virtual scholars working in a rich digital information environment. It also clarified (see section 5.5.1) the issue of use of Google by physicists for finding articles that was raised by a previous research (Nicholas et al, 2005b).

The study is also one of the few studies that applied mixed-methods, especially the combination of qualitative and quantitative approach by using interview, questionnaire and information-event cards for data collection. The study showed that mixed-methods could help collect a richer and comprehensive dataset and at the same time it exposed the challenges of

using mixed-methods (qualitative and quantitative) for investigating information behaviour and highlighted the difficulties it may cause.

This thesis is also one of the few studies that focused on the different aspects of information-seeking behaviour of scientists including the similarities and dissimilarities in information-seeking behaviour of academic scientists based on their academic status, the type of their research, their field of research as well as the stages of a research project.

### **6.7. Models of information-seeking behaviour**

This study was not designed to build a model of information-seeking behaviour. However, the ultimate aim of research conducted in different contexts is thought to be identifying general patterns, modes, or processes of information seeking (Talja, Keso & Pietiläinen, 1999, 759). The results of this study showed some level of similarity in information-seeking behaviour of people in the different subfields of physics and astronomy (for example the popularity of human information resources). However, the study also revealed several differences in their information-seeking behaviour and showed that it is somewhat over-generalising to consider the whole field of physics and astronomy as a single field whose scientists all behave in a very similar way that can be presented in a single model.

The identified intradisciplinary differences make it hard to propose a general model for the information-seeking behaviour of all physicists and astronomers, let alone for all scientists. With some level of compromise some of the past studies produced models of information-seeking behaviour that are meant to be applicable to specific groups of information users or to specific information seeking tasks. Some of these models such as Ellis's (which was discussed in chapter two) are claimed to be applicable to all scientists.

Ellis's model is a very general model with broad categories of information-seeking behaviour (such as Starting and Ending) and that is the reason why it is applicable to users in different academic disciplines. Revisiting Ellis's model shows that all of the eight categories that were found by Ellis, Cox and Hall (1993) in the behaviour of physical scientists have been applied by the subjects of the current study. However, different research groups relied on different means and methods in their information seeking. For example in the case of 'Starting', while some groups (such as tracking TMP) relied mainly on tracking references, some other groups (such as HEP) had more reliance on searching.

In general, while models such as Ellis's model are helpful for knowing about general types of behaviour that scientists may express in their information seeking activities, the results of this study show that a closer look at inter- and intra-disciplinary differences is

necessary for designing and providing information services that could efficiently meet the information needs of scientists at an individual level.

## 6.8. Further research

This study investigated the information-seeking behaviour of physicist and astronomers with an intradisciplinary approach and indicated that although similarities exist among different areas, each subfield has its own characteristics. The study implies the importance of taking a micro approach and looking at smaller subject areas while investigating the information behaviour. Although study of broad disciplines is helpful in our general understanding of information behaviour of scientists, narrower subject areas should be studied if the information services are to move toward more personalisation and specialisation. As scientific fields become more and more specialised, this approach to the study of information behaviour could enhance our understanding of scientists' information behaviour and can lead to more effective and efficient information services and systems. Therefore, adopting an intradisciplinary approach for the study of the information behaviour of scholars in other disciplines could be a possibility for the future work.

Birger Hjørland's (2004) domain-analytic approach could be helpful for this type of research although its application could be challenging and problematic. The researcher did not use this approach in the present study due to a number of reasons including the fact that it is still in its infancy and rather ambiguous, the lack of clear definition for domain, the fact the proposed framework by Hjørland is too broad and disparate to form a blue print for a PhD thesis. However, adopting this approach as it becomes more developed and studying information behaviour of a discipline by looking at subfields within it could be a fruitful study.

This has been the first study in its kind to investigate the information-seeking behaviour of physicists and astronomers with intradisciplinary approach. However, due to the limitations associated with that the population of the study, which is restricted to a single department; it is a good idea to conduct an international survey of physicist and astronomers. Such a study could validate the findings of this study and get a better sense of the intradisciplinary differences and similarities in the information-seeking behaviour of physicists and astronomers in a global scale.

The scope of this study covered several aspects of information-seeking behaviour of physicists and astronomers. Although such a broad scope is helpful in gaining a fuller picture of general information-seeking behaviour of a subject community, it might be more effective to focus on a single aspect of the information behaviour of a community while conducting intradisciplinary studies.

Besides further intradisciplinary investigation of information behaviour of scientists, the study raised some questions and highlighted some issues that could be each a subject for another research study. These have been mentioned in chapter five and this chapter in relevant places; however they are mentioned briefly here again.

One potential area for further research is the nature of reading by scientists. Research needs to be done in order to understand reading as a phenomenon and its nature. We still know little about how scientists read the papers, whether their reading differs when they read for different tasks and purposes and if yes, how. Knowing about these issues is important because it has implications. For example one important outcome that our better understanding of reading could have is related to the critical and increasingly important issue of use of digital resources. At the moment there is no clear definition of what usage actually means, especially in the digital environment. In the case of electronic journals, most research studies so far (Rowlands and Nicholas, 2008) consider a download as a use. However there is no guarantee that a downloaded article is used. As publishers and librarians use the full-text download metric as a basis for the judgement for the value of their service, one must look more closely to this issue of how scientists actually use the resources. These issues could be the subject of a PhD thesis themselves.

Another potential area for further research is the way scientists interact with e-print archives, the disciplinary and intradisciplinary differences and the driving factors of their behaviour. The focus of most of the existing literature is on issues such as the use or non-use of e-print archives (e.g. Lawal, 2002) and the impact of e-print archives and open access resources on scholarly communication (e.g. Harnad and Brody, 2004; Antelman, 2004). The current research showed that although generally physicists and astronomers are the highest users of e-print archives (see also Kling & McKim, 2000; Fry, 2003), there are variations within physics and astronomy not only in the amount of use of e-print archives (both in terms of depositing papers and reading papers), but also in terms of the way they interact with e-print archives. Clarifying these variations by conducting a research study which only focuses on the interaction of physicists and astronomers with e-print archives and investigating the driving factors in the behavioural variations of their interaction with the archive is a needed research.

## 6.9. A final word

Although the limitations of the study were discussed in section 3.9, we should be reminded here that one must be cautious in generalisation of the findings of the study. Like many other user studies in the field of library and information sciences this study was restricted to a sample drawn from a specific organisation and as the results themselves showed the

organisational factors (such as research environment) affect information-seeking behaviour of academics.

As this thesis was being written the Department of Physics and astronomy went through some changes in 2007. A new building was built for the London Centre for Nanotechnology that now accommodates the CMMP group; and some of the other groups such Atmospheric Physics that were located in different places inside and outside the main campus at Gower Street moved to the main building of the Department. These physical movements might have affected the research environment and the dynamics of the communications within each group and among different groups, and that in turn would affect the information-seeking behaviour of the people in the department.

Besides these physical movements, another issue raises questions about generalisability of the findings of this study about specific research areas of physics and astronomy at UCL to those areas generally in physics and astronomy. This issue is the change in the Department's structure. Some of the smaller research groups merged and new research groups were shaped. For example the Astronomy and Astrophysics was restructured. This dynamic of academic research domains shows how complex the area of scholarly communication and information-seeking behaviour is when it comes to inter- and intra-disciplinary comparisons.

One last issue worth paying attention to is the fact that the studied department is a large and prestigious one. One might argue that large departments are typical departments in which hard science is done therefore it is not irrational to generalise the findings about this department to physicists and astronomers. On the other hand, it should be appreciated that the size of the department might affect information-seeking behaviour of the academics and for example physicists who work in a smaller department might behave differently in terms of their information seeking. For example, when there are fewer colleagues the reliance on colleagues as sources of information might reduce hence more active measures might be taken by users to satisfy their needs while in large departments with large research groups there might expectedly be more opportunities for collaborative information seeking and sharing information.

## 6.10. Summary

The chapter drew the conclusions and discussed them in the light of past studies. It also presented some suggestions for further investigations. The chapter illustrated the main characteristics of the information-seeking behaviour and scholarly communication of the different subfields of physics and astronomy. It showed that some differences exist within physics and astronomy in terms of different aspects of information seeking. The chapter also



showed the similarities and dissimilarities among physicists and astronomers with different status with regard to their information-seeking behaviour and scholarly communication. Moreover, the chapter described the information-seeking behaviour of physicists and astronomers in the different stages of a research project and explained the main characteristics of the information seeking behaviour in each stage of a research project. The key characteristics of reading behaviour of physicists and astronomers as well as their publishing behaviour in terms of their interaction with journals and e-print archives were also discussed.

## BIBLIOGRAPHY

- Abdoulaye, K. (2002). Information-Seeking Behaviour of African Students in Malaysia: a research study. *Information Development*, **18** (3), 191-196.
- Allan, J., Carterette, B., & Lewis, J. (2005). "When Will Information Retrieval Be "Good Enough"?" (pp. 433-440). *In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, Salvador, Brazil. ACM Press.
- Allen, B.L. (1991). "Cognitive research in information science: implications for design". *In* M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 26, pp. 3-37). Medford, NJ: Learned Information.
- Allen, B.L. (1996). *Toward a user-centered approach to information systems*. San Diego, CA: Academic Press.
- Allen, R.S. (1991). "Physics information and scientific communication: information sources and communication patterns". *In* C. Steinke (Ed.), *Information seeking and communicating behavior of scientists and engineers* (pp. 27-38). New York: Haworth Press.
- Allen, T.J. (1969). "Information needs and uses". *In* C.A. Cuadra (Ed.), *Annual Review of Information Science and Technology* (Vol. 4, pp. 3-30). Chicago: Encyclopaedia Britannica.
- Allen, T.J. (1977). *Managing the flow of information: technology transfer and the dissemination of technical information within the R & D organization*. Cambridge, MA: MIT Press.
- Antelman, K. (2004). Do Open-Access articles have a greater research impact? *College and Research Libraries*, **65** (5), 372-382.
- arXiv. (2007). arXiv monthly submission rate statistics, from [http://arxiv.org/show\\_monthly\\_submissions](http://arxiv.org/show_monthly_submissions) (visited 25/11/2007).
- Baldwin, N.S., & Rice, R.E. (1997). Information-seeking behavior of securities analysts: Individual and institutional influences, information sources and channels, and outcomes.

- Journal of the American Society for Information Science*, **48** (8), 674-693.
- Banwell, L., & Gannon-Leary, P. (2000). JUBILEE: monitoring user information behaviour in the electronic age. *OCLC Systems & Services*, **16** (4), 189-193.
- Barrett, A. (2005). The Information-Seeking Habits of Graduate Student Researchers in the Humanities. *Journal of Academic Librarianship*, **31** (4), 324-331.
- Barry, C.A. (1995). Critical Issues in Evaluating the Impact of IT on Information Activity in Academic Research: Developing a Qualitative Research Solution. *Library and Information Science Research*, **17** (2), 107-134.
- Barry, C.A. (1997). Information Skills for an Electronic World: Training Doctoral Research Students. *Journal of Information Science*, **23** (3), 225-238.
- Bates, M.J. (1977). Factors Affecting Subject Catalog Search Success. *Journal of the American Society for Information Science*, **28** (3), 161-169.
- Bates, M.J. (1996). Learning about the information seeking of interdisciplinary scholars and students. *Library Trends*, **45** (2), 155-164.
- Bates, M.J. (2002). "Speculations on browsing, directed searching, and linking in relation to the Bradford distribution". In H. Bruce, R. Fidel, P. Ingwersen & P. Vakkari (Eds.), *Emerging Frameworks and Methods: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS 4)*, July 2 1-25 (pp. 137-149). Seattle, WA: Libraries Unlimited, Greenwood Village.
- Bawden, D. (2001). *Information overload*. London: South Bank University, Library Information Technology Centre.
- Bawden, D. (2006). Users, user studies and human information behaviour: A three-decade perspective on Tom Wilson's "On user studies and information needs". *Journal of Documentation*, **62** (6), 671-679.
- Bawden, D., & Vilar, P. (2006). Digital libraries: to meet or manage user expectations. *Aslib Proceedings*, **58** (4), 346-354.
- Bazerman, C. (1985). Physicists reading physics: schema-laden purposes and purpose-laden schema. *Written Communication*, **2** (1), 3-23.
- Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the*

*Experimental Article in Science*. Madison, WI: The University of Wisconsin Press.

Becher, T., & Trowler, P.R. (2001). *Academic tribes and territories: intellectual enquiry and the culture of disciplines* (2nd ed.). Buckingham: Open University Press.

Beghtol, C. (1995). Domain analysis, literary warrant, and consensus: The case of fiction studies. *Journal of the American Society for Information Science*, **46** (1), 30-44.

Belkin, N.J. (1990). The cognitive viewpoint in information science. *Journal of Information Science*, **18** (1), 11-15.

Bernstein, B. (2000). *Pedagogy, symbolic control and identity: theory, research, critique*. Lanham, MD: Rowan & Littlefield.

Bigdeli, Z. (2007). Iranian engineers' information needs and seeking habits: an agro-industry company experience. *Information Research*, **12** (2), paper 290. <http://informationr.net/ir/12-2/paper290.html> (visited 25/11/2007).

Brannen, J. (2004). "Working qualitatively and quantitatively". In C. Seale, G. Gobo, J.F. Gubrium & D. Silverman (Eds.), *Qualitative research practice* (pp. 312-326). London: Sage.

Brannen, J., & Moss, P. (1991). *Managing mothers: dual earner households after maternity leave*. London: Unwin Hyman.

Brockman, W.S., Neumann, L., Palmer, C.L., & Tidline, T.J. (2001). *Scholarly Work in the Humanities and the Evolving Information Environment*. Washington, DC: Digital Library Federation, Council on Library and Information Resources.

Brown, C.M. (1999a). Information literacy of physical science graduate students in the information age. *College and Research Libraries*, **60** (5), 426-438.

Brown, C.M. (1999b). Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science*, **50** (10), 929-943.

Brown, C.M. (2001a). The Coming of Age of E-Prints in the Literature of Physics. *Issues in Science and Technology Librarianship*, **31** (Summer). <http://www.istl.org/01-summer/refereed.html> (visited 25/11/2007).

Brown, C.M. (2001b). The E-evolution of Preprints in the Scholarly Communication of

- Physicists and Astronomers. *Journal of the American Society for Information Science and Technology*, **52** (3), 187-200.
- Bryman, A. (2001). *Social Research Methods*. Oxford: Oxford University Press.
- Budd, J.M., & Connaway, L.S. (1997). University faculty and networked information: results of a survey. *Journal of the American Society for Information Science*, **48** (9), 843–852.
- Bureau of Labor Statistics, U.S. Department of Labor. (2006). *Occupational Outlook Handbook, 2006-07 Edition, Physicists and Astronomers*. U.S. Department of Labor. (2006-07 ed.).
- Callinan, J.E. (2005). Information-seeking behaviour of undergraduate biology students: A comparative analysis of first year and final year students in University College Dublin. *Library Review*, **54** (2), 86 - 99.
- Case, D.O. (2002). *Looking for information: a survey of research on information seeking, needs, and behaviour*. Amsterdam: Academic Press.
- Case, D.O. (2005). "Information behavior". In B. Cronin (Ed.), *Annual Review of Information Science and Technology* (Vol. 40). Medford, NJ: Information Today.
- Chang, N.C., & Perng, J.H. (2001). Information Search Habits of Graduate Students at Tatung University. *The International Information & Library Review*, **33** (4), 341-346.
- Chatman, E.A. (1985). Information, mass media use and the working poor. *Library and Information Science Research*, **7** (1), 97-113.
- Chatman, E.A. (1990). Alienation theory: application of a conceptual framework to a study of information among janitors. *RQ*, **29** (3), 355-368.
- Chatman, E.A. (1991a). Channels to a larger social world: older women staying in contact with the great society. *Library and Information Science Research*, **13** (3), 281-300.
- Chatman, E.A. (1991b). Life in a small world: applicability of gratification theory to information-seeking behavior. *Journal of the American Society for Information Science*, **42** (6), 438-449.
- Chatman, E.A. (1999). A theory of life in the round. *Journal of the American Society for Information Science*, **50** (3), 207-217.
- Chatman, E.A. (2000). Framing social life in theory and research. *The New Review of*

*Information Behaviour Research: Studies of Information Seeking in Context*, 1, 3-17.

- Cho, A. (2000). Distorted Galaxies Point to Dark Matter. *Science*, 287 (5460), 1899-1901.
- Chu, S. (1995). "Difficulties that may be Encountered by Students in Searching the Dow Jones News/ Retrieval Database and the UMI Proquest Fullimage Databases" (Vol. 19, pp. 81-100). *Proceedings of the International Online Information Meeting*, London, UK.
- Chu, S.K.-W., & Law, N. (2005). Development of information search expertise: research students' knowledge of databases. *Online Information Review*, 29 (6), 621-642.
- Chu, S.K.-W., & Law, N. (2007). Development of information search expertise: research students' knowledge of source types. *Journal of Librarianship and Information Science*, 39 (1), 27-40.
- Cole, C. (2000). Inducing Expertise in History Doctoral Students via Information Retrieval Design. *Library Quarterly*, 70 (1), 86-109.
- Compton, M.L. (1989). *A Study of the Information Resources and Library Services used by Doctoral Students in science Education at the University of Georgia*. Unpublished Masters thesis, University of Georgia.
- Cornelius, I. (2002). "Theorizing information for information science". In B. Cronin (Ed.), *Annual Review of Information Science and Technology* (Vol. 36, pp. 392-425). Medford, NJ: Information Today.
- Couper, M.P. (2000). Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly*, 64 (4), 464-494.
- Crane, D. (1971). "Information needs and uses". In C.A. Cuadra (Ed.), *Annual Review of Information Science and Technology* (Vol. 6, pp. 3-39). Chicago: Encyclopaedia Britannica.
- Crane, D. (1972). *Invisible colleges: diffusion of knowledge in scientific communication*. Chicago: University of Chicago Press.
- Crawford, S. (1978). "Information needs and uses". In M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 13, pp. 61-84). White Plains, NY: Knowledge Industry Publications.
- Crawford, S., McCabe, S., Couper, M., & Boyd, C. (2002). "From mail to Web: improving

- response rates and data collection efficiencies". Paper presented at *the International Conference on Improving Surveys*, 25-28 August, University of Copenhagen - Denmark.
- Creswell, H.W. (1995). *Research design: qualitative and quantitative approaches*. Thousands Oaks, CA: Sage.
- Cronin, B. (1982). Progress in documentation: invisible colleges and information transfer, a review and commentary with particular reference to the social sciences. *Journal of Documentation*, **38** (3), 212-236.
- De May, M. (1977). "The cognitive viewpoint: its development and its scope". In M. De Mey, R. Pinxten, M. Poriau & F. Vandamme (Eds.), *International Workshop on the Cognitive Viewpoint* (pp. xvi-xxxii). Ghent: University of Ghent.
- Denzin, N.K., & Lincoln, Y.S. (1998). *Collecting and Interpreting Qualitative Materials*. London: Sage.
- Dervin, B. (1983). *An overview of sense-making: concepts, methods, and results to date, International Communication Association Annual Meeting*. Dallas, Texas.
- Dervin, B., & Nilan, M. (1986). "Information needs and uses". In M.E. Williams (Ed.), *Annual review of information science and technology* (Vol. 21, pp. 3-33). White Plains, NY: Knowledge Industry Publications.
- Dixon, A. (1999). The twelve ages of electronic journals. *Vine* (111), 3-9.
- Drabenstott, K.M., & Weller, M.S. (1996). Failure Analysis of Subject Searches in a Test of a New Design for Subject Access to Online Catalogs. *Journal of the American Society for Information Science*, **47** (7), 519-537.
- Edwards, P., Roberts, I., Clarke, M., DiGuseppi, C., Prata, S., Wentz, R., et al. (2002). Increasing response rates to postal questionnaires: systematic review. *British Medical Journal*, **324** (7347), 1183.
- Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, **45** (3), 171-212.
- Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, **49** (4), 356-369.

- Ellis, D., & Haugan, M. (1997). Modelling the information-seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, **53** (4), 384-403.
- Fabiano, E. (1996). Casting the Net: Reaching out to Doctoral Students in Education. *Research Strategies*, **14** (3), 159-168.
- Farid, M.A. (1984). *A Study of Information Seeking Behavior of PhD Students in Selected Disciplines. Final Report*. New York: Syracuse University, School of Information Studies.
- Fidzani, B.T. (1998). Information needs and information-seeking behaviour of graduate students at the University of Botswana. *Library Review*, **47** (7), 329-340.
- Fisher, S., & Oulton, T. (1999). The Critical Incident Technique in Library and Information Management Research. *Education for Information*, **17** (2), 113-125.
- Flanagan, J.C. (1954). The Critical Incident Technique. *Psychological Bulletin*, **51** (4), 327-358.
- Fluckiger, F. (1989). Overview of HEP wide networking; producer perspective. *Computer Physics Communication*, **57** (11), 183-187.
- Fogg, B.J. (1999). Persuasive technologies - Now is your chance to decide what they will persuade us to do - and how they'll do it. *Communications of the ACM*, **42** (1), 26-29.
- Ford, N., & Mansourian, Y. (2006). The invisible web: An empirical study of 'cognitive invisibility'. *Journal of Documentation*, **62** (5), 584-596.
- Fox, J.W. (200?) *From Lander to Massey: History of physics, space science and astronomy at University College London, 1826 - 1975*.  
<http://www.phys.ucl.ac.uk/department/history/BFox1.html> (visited 25/11/2007).
- Frank, E. (1994). Authors' criteria for selecting journals. *JAMA: The Journal of the American Medical Association*, **272**, 163-164.
- Frankel, M.S., Elliott, R., Blume, M., Bourgois, J.-M., Hugenholtz, B., Lindquist, M.G., et al. (2000). Defining and Certifying Electronic Publication in Science: A Proposal to the International Association of STM Publishers. *Learned Publishing*, **13** (4), 251-258.
- Froehlich, T.J. (1994). Relevance reconsidered-Towards an agenda for the 21st century:



- Introduction to special topic issue on relevance research. *Journal of the American Society for Information Science*, **45** (3), 124-134.
- Fry, J. (2003). *The cultural shaping of scholarly communication within academic specialisms*. Unpublished PhD Thesis, University of Brighton, Brighton.
- Fry, J. (2006). Scholarly research and information practices: a domain analytic approach. *Information Processing and Management*, **42** (1), 299-316.
- Fry, J., & Talja, S. (2004). "The cultural shaping of scholarly communication: explaining e-journal use within and across academic fields". Paper presented at *the American society for information science and technology annual meeting on managing and enhancing information: cultures and conflicts*. , 13-18 November, Providence, Rhode Island.
- Fry, J., & Thelwall, M. (2006). *Using Domain Analysis and Organisational Theory to Understand E- Science Sustainability*. Paper presented at the Second International Conference on e-Social Science, 28-30 June 2006, Manchester, UK.
- Garvey, W.D. (1979). *Communication: the essence of science*. Oxford: Pergamon Press.
- George, C., A. Bright, Hurlbert, T. et al. (2006). Scholarly use of information: graduate students' information seeking behaviour. *Information Research*, **11** (4), paper 272. <http://informationr.net/ir/11-4/paper272.html> (visited 10/02/2008).
- Gerstberger, P.G., & Allen, T.J. (1968). Criteria used by research and development engineers in the selection of an information source. *Journal of applied psychology*, **52** (4), 272-279.
- Ginsparg, P. (1994). First steps towards electronic research communication. *Computers in Physics*, **8** (4), 390-396.
- Gobo, G. (2004). "Sampling, representativeness and generalizability". In C. Seale, G. Gobo, J.F. Gubrium & D. Silverman (Eds.), *Qualitative research practice* (pp. 435-456). London: Sage.
- Gordon, M.D. (1984). How authors select journals: a test of the reward maximization model of submission behaviour. *Social Studies of Science*, **14** (1), 27-43.
- Gorman, G.E., & Clayton, P. (2005). *Qualitative research for the information professional: a practical handbook* (2nd ed. ed.). London: Facet.

- Gould, C. C. and Pearce, K. (1991). *Information Needs in the Sciences: an Assessment*. Mountain View, California: Research Library Group.
- Greene, S., Marchionini, G., Plaisant, C., & Shneiderman, B. (2000). Previews and overviews in digital libraries: designing surrogates to support visual information seeking. *Journal of the American Society for Information Science*, **51** (4), 380-393.
- Gresham, J.L. (1994). From invisible college to cyberspace college: computer conferencing and the transformation of informal scholarly communication networks. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century*, **2** (4), 37-52.
- Guba, E.G., & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- Gunter, B., Nicholas, D., Huntington, P., & Williams, P. (2003). Digital interactive television: health information platform of the future? *Aslib Proceedings*, **55** (5-6), 346-356.
- Hagstrom, W.O. (1970). "Factors related to use of different modes of publishing research in four scientific fields". In C.E. Nelson & D.K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 85-124). Lexington, Massachusetts: Heath Lexington Book.
- Hajjem, C., Harnad, S., & Gingras, Y. (2005). Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *IEEE Data Engineering Bulletin*, **28** (4), 39-47. <http://eprints.ecs.soton.ac.uk/12906/> (visited 25/11/2007)
- Hallmark, J. (2004). Access and Retrieval of Recent Journal Articles: A Comparative Study of Chemists and Geoscientists. *Issues in Science and Technology Librarianship* **40**. <http://www.istl.org/04-summer/article1.html> (visited 25/11/2007).
- Hammersley, M. (1987). Some notes on the terms 'validity' and 'reliability'. *British Educational Research Journal*, **13** (1), 73-81.
- Hammersley, M. (1992). "Deconstructing the qualitative- quantitative divide". In J. Brannen (Ed.), *Mixing methods: qualitative and quantitative research* (pp. 39-57). London: Gower.
- Hammersley, M. (1996). "The relationship between qualitative and quantitative research: paradigm loyalty versus methodological eclecticism". In J.T.E. Richardson (Ed.), *Handbook of research methods for psychology and the social sciences* (pp. 159-179).

Leicester: BSP Books.

Harnad, S., & Brody, T. (2004). Comparing the impact of Open Access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, **10** (6). <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/june04/harnad/06harnad.html> (visited 25/11/2007).

Hemminger, B.M., Lu, D., Vaughan, K.T.L., & Adams, S.J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, **54** (14), 2205-2225.

Hensman, S. (1977). *Conference attendance patterns amongst physicists*. London: Aslib.

Herner, S. (1954). Information gathering habits of workers in pure and applied science. *Industrial and Engineering Chemistry*, **46** (1), 228-236.

Herner, S., & Herner, M. (1967). "Information needs and uses in science and technology". In C.A. Cuadra (Ed.), *Annual Review of Information Science and Technology* (Vol. 2, pp. 1-34). New York: Interscience Publishers.

Hertzum, M., & Pejtersen, A.M. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing and Management*, **36** (5), 761-778

Hess, B. (1999). Graduate Student Cognition during information Retrieval using the World Wide Web: A Pilot Study. *Computers & Education*, **33** (1), 1-13.

Hewins, E.T. (1990). "Information need and use studies". In M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 25, pp. 145-172). Amsterdam: Elsevier Science.

Hildreth, C.R. (1997). The Use and Understanding of Keyword Searching in a University Online Catalog. *Information Technology and Libraries*, **16** (2), 52-62.

Hjørland, B. (1993). "Toward a new horizon in information science (I.S.): Domain-analysis", Paper presented at *ASIS 56's Annual Meeting*. Columbus, Ohio, 25 October 1993.

Hjørland, B. (2002a). Domain analysis in information science: Eleven approaches – traditional as well as innovative. *Journal of Documentation*, **58** (4), 422-462.

Hjørland, B. (2002b). Epistemology and the socio-cognitive perspective in information

- science. *Journal of the American Society for Information Science and Technology*, **53** (4), 257-270.
- Hjørland, B. (2004). "Domain analysis in information science". In A. Kent (Ed.), *Encyclopedia of Library and Information science* (pp. 1-7). New York: Marcel Dekker
- Hjørland, B. (2007). *Domain analysis*. Email to Hamid R. Jamali on 20 February 2007.
- Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, **46** (6), 400-425.
- Hjørland, B., & Hartel, J. (2003). Afterword: ontological, epistemological and sociological dimensions of domains. *Knowledge Organization*, **30** (3-4), 239-245.
- Holland, M.P., Pinelli, T.E., Barclay, R.O., & J.M., K. (1991). NASA/DoD Aerospace Knowledge Diffusion Research Project. *European Journal of Engineering Education*, **16** (4), 317-336.
- Huntington, P., Nicholas, D., Jamali, H.R., & Tenopir, C. (2006). Article decay in the digital environment: a usage analysis by date of publication employing deep log methods. *Journal of the American Society for Information Science and Technology*, **57** (13), 1840-1851.
- Hurd, J. M. (1996). "High energy physics". In S. Crawford, J.M. Hurd & A.C. Weller, (Eds.), *From print to electronic: The transformation of scientific communication* (pp. 65-76). Medford, NJ: Information Today.
- Ingwersen, P. (1984). Psychological aspects of information retrieval. *Social Science Information Studies*, **4** (2-3), 83-89.
- Ingwersen, P. (1995). "Information and information science". In K. Allen (Ed.), *Encyclopedia of Information Science* (Vol. 56, Supplement 19, pp. 137-174). New York: Marcel Dekker.
- Ingwersen, P. (1996). Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation*, **52** (1), 3-50.
- Ingwersen, P., & Järvelin, K. (2005). *The turn: integration of information seeking and retrieval in context*. Dordrecht: Springer.

- Jamali, H.R., Nicholas, D., & Huntington, P. (2005). The use and users of scholarly e-journals: a review of log analysis studies. *Aslib Proceedings*, 57 (6), 554-571.
- Jankowska, M.A., Hertel, K., & Young, N.J. (2006). Improving Library Service Quality to Graduate Students: LibQual+™ Survey Results in a Practical Setting. *Portal: Libraries and the Academy*, 6 (1), 59-77.
- Jastrow, D. (2004). *Global STM Market: Analysis and Forecast 2004*. Stamford: Simba Information.
- Johnson, B., & Christensen, L. (2000). *Educational research: qualitative and quantitative approaches*. Boston: Allyn & Bacon.
- Kaplowitz, M.D., Hadlock, T.D., & Levine, R. (2004). A Comparison of Web and Mail Survey Response Rates. *Public Opinion Quarterly*, 68 (1), 94-101.
- Kasperson, C.J. (1978). Scientific creativity: a relationship with information channels. *Psychological Reports*, 42 (3), 691-694.
- Katzer, J., & Fletcher, P.T. (1992). "The information environment of managers". In M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 27, pp. 227-263). Medford, NJ: Learned Information.
- King, D.W., McDonald, D.D. & Roderer, N.K. (1981). *Scientific Journals in the United States: Their Production, Use and Economics*. Stroudsburg, PA: Hutchinson Ross Publishing Company (Division of Academic Press).
- King, D.W., & Montgomery, C.H. (2002). After Migration to an Electronic Journal Collection: Impact on Faculty and Doctoral Students. *D-Lib Magazine*, 8 (12). <http://www.dlib.org/dlib/december02/king/12king.html> (visited 25/11/2007)
- King, D.W., & Roderer, N.K. (1982). Communication in physics - the use of journals. *Physics Today* (October), 44-47.
- King, D.W., & Tenopir, C. (2001). "Using and reading scholarly literature". In M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 34, pp. 423-477). Medford, NJ: Information Today.
- Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences in the shaping of electronic media in supporting scientific communication. *Journal of the American*

*Society for Information Science and Technology*, **51** (14), 1306-1320.

Knorr Cetina, K.D. (1999). *Epistemic cultures: how the sciences make knowledge*. London: Harvard University Press.

Kraut, R.E., Egido, C., & Galegher, J. (1990). "Patterns of contact and communication in scientific research collaboration". In J. Galegher, R.E. Kraut & C. Egido (Eds.), *Intellectual teamwork: Social and technological foundations of cooperative work* (pp. 149-171). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Krikelas, J. (1983). Information-Seeking Behavior: Patterns and Concepts. *Drexel Library Quarterly*, **19** (6), 5-20.

Krueger, R.A. (1994). *Focus group: a practical guide for applied research* (2nd ed.). Thousands Oaks, CA: SAGE.

Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science and Technology*, **42** (5), 361-371.

Kurata, K., Matsubayashi, M., Mine, S., Muranushi, T., & Ueda, S. (2007). Electronic journals and their unbundled functions in scholarly communication: Views and utilization by scientific, technological and medical researchers in Japan. *Information Processing and Management*, **43** (5), 1402-1415.

Lancaster, F. W. (2003). *Indexing and abstracting in theory and practice* (3rd ed.). London: Facet.

Lawal, I. (2002). Scholarly Communication: The Use and Non-Use of E-Print Archives for the Dissemination of Scientific Information. *Issues in Science and Technology Librarianship* (36). <http://www.istl.org/02-fall/index.html> (visited 25/11/2007).

Leckie, G.J., Pettigrew, K.E., & Sylvain, C. (1996). Modelling the information seeking of professionals: A general model derived from research on engineers, health professionals and lawyers. *The Library Quarterly*, **66** (2), 161-193.

Lee, A.S., & Baskerville, R.L. (2003). Generalizing Generalizability in Information Systems Research. *Information Systems Research*, **14** (3), 221-243.

Lewis, J., & Ritchie, J. (2003). "Generalising from qualitative research". In J. Ritchie & J.

- Lewis (Eds.), *Qualitative research practice: a guide for social science students and researchers* (pp. 263-286). London: Sage.
- Leydesdorff, L., & Zhou, P. (2005). Are the contributions of China and Korea upsetting the world system of science? *Scientometrics*, **63** (3), 617-630.
- Lindström, P. (2005). Vaiheista sykleihin: tiedonhankinnan uusia malleja. *Informaatiotutkimus*, **24** (1), 211.
- Liull, H. (1991). Meeting the academic research information needs of the scientists and engineers in the university environment. *Science & Technology Libraries*, **11** (3), 83-90.
- Lynch, C. (2007). The Shape of the Scientific Article in the Developing Cyberinfrastructure. *CTWatch Quarterly*, **3** (3). <http://www.ctwatch.org/quarterly/articles/2007/08/the-shape-of-the-scientific-article-in-the-developing-cyberinfrastructure/> (visited 25/11/2007).
- Mabe, M. (2003). The growth and number of journals. *Serials: The Journal for the Serials Community*, **16** (2), 191-197.
- Majid, S., & Tan, A.T. (2002). Usage of information resources by computer engineering students: a case study of Nanyang Technological University, Singapore. *Online Information Review*, **26** (5), 318-325.
- Makri, S., Blandford, A., & Cox, A.L. (2008). Investigating the information-seeking behaviour of academic lawyers: From Ellis's model to design. *Information Processing and Management*, **44** (2), 613-634.
- Mansourian, Y. (2006). *Information visibility on the Web and conceptions of success and failure in Web searching*. Unpublished PhD, The University of Sheffield, Sheffield.
- Marchionini, G. (1995). *Information seeking in electronic environment*. Cambridge: Cambridge University Press.
- McCoy, S.V., & Marks, P. (2001). "Using electronic surveys to collect data: experiences from the field" (pp. 1502-1505). Paper presented at *the Seventh Americas Conference on Information Systems*, 3-5 August, Bentley College, Boston.
- Meho, L. (2006). E-mail interviewing in qualitative research: A methodological discussion. *Journal of the American Society for Information Science and Technology*, **57** (10),

1284-1295.

- Meho, L.I., & Tibbo, H.R. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology*, **54** (6), 570-587.
- Menzel, H. (1964). The information needs of current scientific research. *The Library Quarterly*, **34** (1), 4-19.
- Metzger, M.J. (2007). Making sense of credibility on the Web: models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, **58** (13), 2078-2091.
- Miles, M., & Huberman, A. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods*. London: Sage.
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed. ed.). London: Sage Publications.
- Miller, D.C. (1991). *Handbook of research design and social measurement* (5th ed.). Newbury Park, Calif.: Sage.
- Mishler, E.G. (1990). Validation in enquiry-guided research: The role of exemplars in narrative studies. *Harvard Educational Review*, **60**, 415-442.
- Morner, C.J. (1993). *An Instrument for Measuring Library Skills of Education Graduate Students*. Unpublished PhD dissertation, Boston College, Boston.
- Morrison, E.W., & Vancouver, J.B. (2000). Within-Person Analysis of Information Seeking: The Effects of Perceived Costs and Benefits. *Journal of Management*, **26** (1), 119-137.
- Moss, J., & Hendry, G. (2002). Use of electronic surveys in course evaluation. *British Journal of Educational Technology*, **33** (5), 583-592.
- National Research Council (1973). *Physics in Perspective* (Vol. 2b). Washington: National Academy of Sciences.
- Neighbors, J. M. (1980). *Software construction using components*. Irvine, CA: University of California. PhD Thesis.
- Nelson, D. (2001). The uptake of electronic journals by academics in the UK, their attitudes towards them and their potential impact on scholarly communication. *Information*



- Services & Use*, **21** (3-4), 205-214.
- Nicholas, D. (2000). *Assessing information needs: tools, techniques and concepts for the internet age*. London: Aslib.
- Nicholas, D., Huntington, P., Dobrowolski, T., Rowlands, I., Jamali, H.R., & Polydoratou, P. (2005a). Revisiting 'obsolescence' and journal article 'decay' through usage data: an analysis of digital journal use by year of publication. *Information Processing & Management*, **41** (6), 1441-1461.
- Nicholas, D., Huntington, P., & Jamali, H.R. (2006). *Authors as users: A deep log analysis linking demographic and attitudinal data obtained from scholarly authors with their usage of ScienceDirect*. London: University College London, CIBER.
- Nicholas, D., Huntington, P., & Jamali, H.R. (2007). The use, users and role of abstracts in the digital scholarly environment. *Journal of Academic Librarianship*, **33** (4), 446-453.
- Nicholas, D., Huntington, P., Rowlands, I., Jamali, H.R., & Haynes, J. (2005b). *An investigation of physicists' information-seeking behaviour, publishing practices, and views on the scholarly publishing system: with special reference to the Institute of Physics and its journals*, Report of an international survey. London: CIBER.
- Nicholas, D., Jamali, H.R., Huntington, P., & Rowlands, I. (2005c). In their very own words: authors and scholarly journal publishing. *Learned Publishing*, **18** (3), 212-220.
- Nicholas, D., Jamali, H.R., & Rowlands, I. (2006). On the tips of their tongues: authors and their views on scholarly publishing. *Learned Publishing*, **19** (3), 193-203.
- Niedźwiedzka, B. (2003). A proposed general model of information behaviour. *Information Research*, **9** (1), Paper 194. <http://InformationR.net/ir/9-1/paper164.html> (visited 25/11/2007).
- Olson, G.M., & Olson, J.S. (2000). Distance matters. *Human-Computer Interaction*, **15** (2/3), 139-178.
- O'Reilly, C.A. (1982). Decision Makers' Use of Information Sources: The Impact of Quality and Accessibility of Information. *The Academy of Management Journal*, **25** (4), 756-771.
- Onwuegbuzie, A.J., Jiao, Q.G., & Bostick, S.L. (2004). *Library Anxiety: theory, research, and*

- applications*. Lanham: The Scarecrow press.
- Oppenheim, A.N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement* (New ed.). London: Continuum.
- Pace, S. (2003). *Understanding the Flow Experiences of Web Users*. Unpublished PhD, Australian National University.
- Pace, S. (2004). A grounded theory of the flow experiences of Web users. *International Journal of Human-Computer Studies*, **60** (3), 327-363.
- Paisley, W.J. (1968). "Information needs and uses". *In* C.A. Cuadra (Ed.), *Annual Review of Information Science and Technology* (Vol. 3, pp. 1-30). Chicago: Encyclopaedia Britannica.
- Palmer, C.L. (1999). Aligning studies of information seeking and use with domain analysis. *Journal of the American Society for Information Science*, **50** (12), 1139-1131 1140.
- Pálsdóttir, Á. (2005). *Health and lifestyle: Icelanders' everyday life information behavior*. Åbo, Finland: Åbo Akademi University Press.
- Pettigrew, K., Fidel, R., & Bruce, H. (2001). "Conceptual frameworks in information behavior". *In* M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 35, pp. 43-78). Medford, NJ: Information Today.
- Pinelli, T.E., Bishop, A.P., Barclay, R.O., & Kennedy, J.M. (1993). "The information seeking behavior of engineers ". *In* *Encyclopedia of Library and Information Science* (Vol. 52 (15), pp. 167-201).
- Pinto, M., & Lancaster, F.M. (1999). Abstracts and Abstracting in Knowledge Discovery. *Library Trends*, **48** (1), 234-248.
- Poland, B.D. (2003). "Transcription quality". *In* J.A. Holstein & J.F. Gubrium (Eds.), *Inside interviewing: new lenses, new concerns* (pp. 267-288). Thousand Oaks, CA: Sage.
- Polydoratou, P., & Moyle, M. (2007). Repository Interface for Overlaid Journal Archives (RIOJA): a feasibility study reporting the views of scientists in astrophysics and cosmology, MTSR 07: 2nd International Conference on Metadata and Semantics Research. Corfu, Greece.
- Poteri, E. (2007). *Collaborative information behaviour: a case study of two research groups*.

Unpublished Licentiate thesis, Åbo Akademi University, Åbo.

Price, D.J.d.S. (1961). *Science since Babylon*. New Haven: Yale University Press.

Punch, K.F. (2005). *Introduction to social research: quantitative and qualitative approaches* (2nd ed.). London: Sage.

Richards, T., & Richards, L. (1998). "Using computers in qualitative research". In N.K. Denzin & Y.S. Lincoln (Eds.), *Collecting and interpreting qualitative materials*. London: Sage.

Rimmer, J., Warwick, C., Blandford, A., Gow, J., & Buchanan, G. (2008). An examination of the physical and the digital qualities of humanities research. *Information Processing and Management*, **44** (3), 1374-1392.

Rocco, T.S., Bliss, L.A., Gallagher, S., & Perez-Prado, A. (2003). Taking the next step: mixed methods research in organisational systems. *Information Technology, Learning, and Performance Journal*, **21** (1), 19-29.

Rowlands, I., & Nicholas, D. (2005). Scholarly communication in the digital environment - The 2005 survey of journal author behaviour and attitudes. *Aslib Proceedings*, **57** (6), 481-497.

Rowlands, I., & Nicholas, D. (2007). The missing link: journal usage metric. *Aslib Proceedings*, **59** (3), 222-228.

Rowley, J. (1994). The controlled versus natural indexing languages debate revisited: a perspective on information retrieval practice and research. *Journal of Information Science*, **20** (2), 108-118.

Royal Society. (1948). Royal Society Scientific Information Conference. London.

Royal Society: Scientific Information Committee (1981). *A study of the scientific information system in the United Kingdom*. London: Royal Society.

Rubenstein, A.H., Buel, W.D., Kegan, D.L., Moore, E.A., Moor, W.C., Rath, G.J., et al. (1970). "Explorations on the information-seeking style of researchers". In C.E. Nelson & D.K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 209-231). Lexington, Mass.: Heath Lexington Books.

Rusch-Feja, D., & Siebeky, U. (1999). Evaluation of Usage and Acceptance of Electronic

- Journals. *D-Lib Magazine*, **5** (10). <http://www.dlib.org/dlib/october99/rusch-feja/10rusch-feja-full-report.html> (visited 25/11/2007).
- Ryen, A. (2004). "Ethical issues". In C. Seale, G. Gobo, J.F. Gubrium & D. Silverman (Eds.), *Qualitative research practice* (pp. 230-247). London: Sage.
- Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, **26**, 321-343.
- Savolainen, R. (2007). Information behavior and information practice: reviewing the 'umbrella concepts' of information-seeking studies. *Library Quarterly*, **77** (2), 109-132.
- Savolainen, R. (2008). Source preferences in the context of seeking problem-specific information. *Information Processing and Management*, **44** (1), 274-293.
- Scott, T.J., & O'Sullivan, M.K. (2005). Analyzing Student Search Strategies: Making a Case for Integrating Information Literacy Skills into the Curriculum. *Teacher Librarian*, **33** (1), 21-26.
- Seale, C.F. (2003). "Computer-assisted analysis of qualitative interview data". In J.A. Holstein & J.F. Gubrium (Eds.), *Inside interviewing: new lenses, new concerns* (pp. 289-308). Thousand Oaks, CA: Sage.
- Self, C.S. (1996). "Credibility". In M. Salwen & D. Stacks (Eds.), *An integrated approach to communication theory and research*. Mahwah, NJ: Erlbaum.
- Shannon, D.M., Johnson, T.E., Searcy, S., & Lott, A. (2002). Using electronic surveys: advice from survey professionals. *Practical Assessment, Research & Evaluation*, **8** (1). <http://PAREonline.net/getvn.asp?v=8&n=> (visited 25/11/2007).
- Shuchman, H.L. (1982). Information technology and the technologist: A report on a national study of American engineers. *International Forum on Information and Documentation*, **7** (1), 3-8.
- Simon, C.E. (1995). *Information Retrieval Techniques: The Differences in Cognitive Strategies and Search Behaviors among Graduate Students in an Academic Library*. Unpublished PhD dissertation, Wayne State University, Michigan.
- Singleton, A. (1997). Journal and the electronic programme of the Institute of Physics. *The*

- Serials Librarian*, 30 (3-4), 149-161.
- Smith, E.T. (2003). Changes in faculty reading behaviours: the impact of electronic journals on the University of Georgia. *Journal of Academic Librarianship*, 29 (3), 162-168.
- Solomon, P. (1997a). Discovering information behavior in sense making. II. The social. *Journal of the American Society for Information Science*, 48 (12), 1109-1126.
- Solomon, P. (1997b). Discovering information behavior in sense making. III. The person. *Journal of the American Society for Information Science*, 48 (12), 1127-1138.
- Sonnenwald, D.H. (1999). "Evolving perspectives of human information behaviour: contexts, situations, social networks and information horizons". In T.D. Wilson & D.K. Allen (Eds.), *Exploring the contexts of information behaviour (Proceedings of The second international conference on research in information needs, seeking and use in different contexts, 13-15 August 1998, Sheffield, UK)*. Sheffield: Taylor Graham.
- Speier, C., Palmer, J., Wren, D., & Hahn, S. (1999). Faculty perceptions of electronic journals as scholarly communication: a question of prestige and legitimacy. *Journal of the American Society for Information Science*, 50 (6), 537-543.
- Spradley, J.P. (1980). *Participant Observation*. New York: Harcourt Brace Jovanovich College Publishers.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks: Sage.
- Sudman, S., & Bradburn, N.M. (1982). *Asking Questions: a Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.
- Talja, S. (2005). "The domain analytic approach to scholar's information practices". In K.E. Fisher, S. Eldelez & L. McKechnie (Eds.), *Theories of information behavior* (pp. 123-127). Medford, NJ: Information Today.
- Talja, S., Keso, H., & Pietiläinen, T. (1999). The production of 'context' in information seeking research: a metatheoretical view. *Information Processing and Management*, 35 (6), 751-763.
- Talja, S., & Maula, H. (2003). Reasons for the use and non use of electronic journals and databases: a domain analytical study in four scholarly disciplines. *Journal of*

*Documentation*, **59** (6), 673-691.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology : combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Taubes, G. (1993). Publication by electronic mail takes physics by storm. *Science*, **259** (5099), 1246-1248.

Tennis, J.T. (2003). Two aces of domains for domain analysis. *Knowledge Organization*, **30** (3-4), 191-195.

Tenopir, C. (2003). Use and users of electronic library resources: a overview and analysis of recent research studies. Report for the Council on Library and Information Resources, from [www.clir.org/pubs/reports/pub120/pub120.pdf](http://www.clir.org/pubs/reports/pub120/pub120.pdf) (visited 25/11/2007).

Tenopir, C., & King, D.W. (1996). Setting the record straight on journal publishing: myth vs. reality. *Library Journal*, **March 15**, 32-35.

Tenopir, C., & King, D.W. (2000). *Towards Electronic Journals: Realities for Scientists, Librarians, and Publishers*. Washington, D.C.: Special Libraries Association.

Tenopir, C., & King, D.W. (2002). Reading behaviour and electronic journals. *Learned Publishing*, **15** (4), 259-265.

Tenopir, C., King, D.W., Boyce, P., Grayson, M., & Paulson, K. (2005). Relying on electronic journals: reading patterns of Astronomers. *Journal of the American Society for Information Science and Technology*, **56** (8), 786-802.

Tenopir, C., King, D.W., Boyce, P., Grayson, M., Zhang, Y., & Ebuon, M. (2003). Patterns of Journal Use by Scientists Through Three Evolutionary Phases. *D-Lib Magazine*, **9** (5). <http://www.dlib.org/dlib/may03/king/05king.html> (visited 25/11/2007).

Tillotson, J. (1995). Is Keyword Searching the Answer? *College and Research Libraries*, **56** (3), 199-206.

Tomney, H., & Burton, P.F. (1998). Electronic journals: a study of usage and attitudes among academics. *Journal of Information Science*, **24** (6), 419-429.

Traweek, S. (1988). *Beamtimes and lifetimes: the world of high-energy physicists*. Cambridge: Harvard University Press.

Tsai, M.J., & Tsai, C.C. (2003). Information Searching Strategies in Web-Based Science

- Learning: The Role of Internet Self-Efficacy. *Innovations in Education and Teaching International*, **40** (1), 43-50.
- Tseng, S., & Fogg, B.J. (1999). Credibility and computing technology. *Communications of the ACM*, **42** (1), 39-44.
- Tuominen, K., Talja, S., & Savolainen, R. (2005). "The Social Constructionist Viewpoint on Information Practices". In K.E. Fisher, S. Eldelez & L. McKechnie (Eds.), *Theories of information behavior* (pp. 328-333). Medford, NJ: Information Today.
- Urquhart, C., Light, A., Thomas, R., Barker, A., Yeoman, A., Cooper, J., et al. (2003). Critical Incident Technique and Explicitation Interviewing in Studies of Information Behavior. *Library & Information Science Research*, **25** (1), 63-88.
- Vakkari, P., & Talja, S. (2005). "The Influence of the scatter of literature on the use of electronic resources across disciplines: a case study of FinELib". In A. Rauber, S. Christodoulakis & A.M. Tjoa (Eds.), *Research and advanced technology for digital libraries: Proceedings of the 9th European Conference, Lecture Notes in Computer Science* (Vol. 3652, pp. 205-217). Berlin: Springer.
- Valauskas, E.J. (1997). Waiting for Thomas Kuhn: First Monday and the evolution of electronic journals. *First Monday*, **2** (12). [http://www.firstmonday.dk/issues/issue2\\_12/valauskas/](http://www.firstmonday.dk/issues/issue2_12/valauskas/) (visited 25/11/2007).
- Walsh, J., & Bayma, P. (1996). Computer networks and scientific work. *Social Studies of Science*, **26** (3), 661-703.
- Wang, P. (2001). "Methodologies and methods for user behavioral research". In M.E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 34, pp. 53-100). Medford, NJ.: Information Today.
- Ware, M. (2005). E-only journals: is it time to drop print? *Learned Publishing*, **18** (3), 193-199.
- Washington-Hoagland, C., & Clougherty, L. (2002). Identifying the Resource and Service Needs of Graduate and Professional Students The University of Iowa User Needs of Graduate Professional Series. *portal: Libraries and the Academy*, **2** (1), 125-143.
- Wathen, C.N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, **53** (2),

134-144.

- Wertman, E.R. (1999). *Electronic preprint distribution: a case study of physicists and chemists at the University of Maryland*. Unpublished MSc, Virginia Polytechnic Institute and State University, Blacksburg, VI.
- Westbrook, L. (1997). "User needs". In A. Kent (Ed.), *Encyclopaedia of Library and Information Science* (Vol. 59, pp. 316-347). New York: Marcel Dekker.
- Whitley, R. (2000). *The intellectual and social organization of the sciences* (2nd ed.). Oxford: Clarendon Press.
- Whitmire, E. (2002). Disciplinary differences and undergraduates' information-seeking behavior. *Journal of the American Society for Information Science and Technology*, **53** (8), 631 - 638.
- Williamson, K., & Bow, A. (2002). "Analysis of quantitative and qualitative data". In K. Williamson (Ed.), *Research methods for students, academics and professionals, information management and systems* (2nd ed., pp. 285-303). Wagga Wagga: Centre for Information Studies, Charles Sturt University.
- Wilson, P. (1993). Communication efficiency in research and development. *Journal of the American Society for Information Science*, **44** (7), 376-382.
- Wilson, T.D. (1981). On user studies and information needs. *Journal of Documentation*, **37** (1), 3-15.
- Wilson, T.D. (1984). The cognitive approach to information-seeking behaviour and information use. *Social Science Information Studies*, **4**, 197-204. <http://informationr.net/tdw/publ/papers/cogapp84.html> (visited 25/11/2007).
- Wilson, T.D. (1999). Models in information behaviour research. *Journal of Documentation*, **55** (3), 249-270.
- Wilson, T.D. (2000). Human information behaviour. *Informing Science*, **3** (2), 49-56. <http://inform.nu/Articles/Vol3/v3n2p49-56.pdf> (visited 25/11/2007).
- Wilson, T.D., & Walsh, C. (1995). *Information behaviour: an interdisciplinary perspective*. London: British Library Research and Development Department.
- Winter, G. (2000). A comparative discussion of the notion of 'validity' in qualitative and



quantitative research. *The Qualitative Report*, 4 (3-4).  
<http://www.nova.edu/ssss/QR/QR4-3/winter.html> (visited 25/11/2007).

Wolcott, H.F. (1990). "On seeking and rejecting validity in qualitative research". In E.W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 121-152). New York: Teachers College Press.

Zaporozhcz, L.E. (1987). *The Dissertation Literature Review: How Faculty Advisors Prepare Their Doctoral Candidates*. Unpublished PhD dissertation, University of Oregon, Oregon.

Zhou, P., & Leydesdorff, L. (2006). The emergence of China as a leading nation in science. *Research Policy*, 35 (1), 83-104.

## APPENDICES

### **Appendix 01: The letter to the Head of the Dept. of Physics and Astronomy**

Monday, 10 October 2005

*To Professor Jonathan Tennyson, Head of the Department of Physics and Astronomy*

#### **Research on the communication and information-seeking behaviour of physicists and astronomers**

Dear Jonathan,

I would like to seek your co-operation for a piece of research that one of my PhD students is undertaking as part of the Virtual Scholar research programme being conducted by my research group, CIBER (<http://www.ucl.ac.uk/ciber/>).

Hamid Jamali's project focuses on the communication and information-seeking behaviour of physicists and astronomers and he would like to conduct a case study at UCL. This research aims to identify communication and information seeking patterns of these groups of scientists; and it attempts to explain the rationales and motivations behind them. The research is particularly interested in the digital information environment.

As part of his investigations he would like to survey postgraduate, research students, researchers and academic staff of your departments. Participation in this project would be voluntary and it would involve an interview of approximately 30 minutes in length to take place in a mutually agreed location. Participants may decline to answer any of the interview questions if they wish. Further, they may decide to withdraw from this project at any time simply by informing the researcher. With their permission, the interview would be tape-recorded and later transcribed for analysis. Their name will not appear in the thesis or any report resulting from this project. All data collected during this project will be kept in a locked and secure place.

I feel confident that this research project would provide interesting results, which will be valuable for the physics and astronomy community. Also, information service providers, such as the UCL library, could benefit from the findings.

If you have any detailed questions regarding this research project, please contact Hamid  
at \_\_\_\_\_ or by email at \_\_\_\_\_

Yours faithfully,

**Professor David Nicholas**

Director, School of Library, Archive and Information Studies

University College London

Tel:

Fax: \_\_\_\_\_

Email: \_\_\_\_\_

Contact Information:

**Hamid R. Jamali**

Research Student,

School of Library, Archive and Information Studies

University College London

Tel:

Fax: \_\_\_\_\_

Email: \_\_\_\_\_

## **Appendix 02: The permission from the Head of the Department of Physics and Astronomy for undertaking the research**

Received on 13th October 2005

Dear both,

Thank you for your letter concerning this research project. I am happy for you to approach members of the Department about this although of course it will be up to individuals at what level they choose to participate. The Department has a weekly electronic newsletter. It might help if you wrote a short piece about this project which we could include in the newsletter.

Best wishes,

Massey Professor of Physics and Head of Department,  
Department of Physics & Astronomy,  
University College London,

Fax:  
Tel:  
Mailto:

**Appendix 03: The note published in the weekly email newsletter of the Dept. of Physics and Astronomy**

UCL DEPARTMENT OF PHYSICS AND ASTRONOMY NEWSLETTER

Monday, 24 Oct 2005

\*\*\*\*\*

RESEARCH ON THE COMMUNICATION AND INFORMATION SEEKING BEHAVIOR OF PHYSICISTS AND ASTRONOMERS

\*\*\*\*\*

Hamid Jamali ( ), a research student at School of Library, Archive and Information Studies and a member of CIBER research group (<http://www.ucl.ac.uk/ciber/ciber.php>) is undertaking research on the communication and information behaviour of physicists and astronomers. The research is part of the project Virtual Scholars which seeks to explore information and communication behaviours of scientist and uncover the rationales and motivations behind them. The research is particularly interested in the digital information environment.

As part of his investigations Hamid would like to survey postgraduate, research students, researchers and academic staff of the department. Participation in this project would be voluntary and it would involve an interview of approximately 30 minutes in length to take place in a mutually agreed location. The individuals will be contacted via email by Hamid for interview and he will be very grateful if you co-operate with him. The research will certainly have interesting results for the physics and astronomy community and also will be useful for improving information services to them.

\*\*\*\*\*

## Appendix 04: The invitation email for participation in the interviews

Dear...

Would it be possible for me to come and talk to you for a half hour about my research? My name is Hamid Jamali, and I'm researching scholarly communication and information-seeking behaviour of physicists and astronomers as part of a PhD thesis in the School of Library, Archive and Information Studies.

Our meeting would only take around half an hour, and would be at a time and location of your choice. Naturally, all information you provide will be completely confidential, and your name will not appear in the thesis or any report resulting from this study. At the same time, I'm happy to give you further details of my work if you wish.

Please note finally, that Professor \_\_\_\_\_ has been informed about this research and permitted me to contact people in your department. I hope that you will have the time to do me this great favour, and would be very grateful if you could contact me at your convenience to arrange a date and time. With many thanks,

Hamid

## Appendix 05: The interview protocol

- Demographics and background: academic background, research description of the interviewee, group membership, type of research the interviewee is mainly involved in.
- How do you approach starting a new research project?
- Are you aware of any patterns or trends in the way research is carried out in your area?
- In what phase of the project do you seek literature? Why? What techniques do you use?
- How do you keep up-to-date with developments in your field of research and related or peripheral fields?
- How do you normally obtain journal articles?
- Discussing use of libraries, digital libraries, repositories and e-print archives and so on.
- Do you have some sorts of criteria to evaluate the information resources you find?
- Describe the process you go through for writing a paper, specifically in terms of information seeking activities it entails?
- How do you decide which journal to submit your paper to? Factors and criteria?
- What is the role of conferences in your information seeking and communication?
- What kind of role do your colleagues and research group play in your information seeking activities?
- Can you think of any problems or difficulties that you have faced in information seeking activities, finding information and locating information resources?
- Have you noticed a notable difference in your information seeking activities and methods from the time you started your doctoral studies till now?

Page. 2	Information-Event Card	No.
4. How much time did you spend looking for this information?		
5. Did you find the information you have been looking for?		
<input type="checkbox"/> A. Yes, found the information needed.		
<input type="checkbox"/> B. Yes, found <u>some</u> of information needed and will continue search.		
<input type="checkbox"/> C. Yes, found <u>some</u> of information needed and will <u>not</u> continue search.		
<input type="checkbox"/> D. No, no information was found.		
<input type="checkbox"/> E. Other (please explain)		
6. For what purpose did you want or need this information?		
7. Describe any unusual events in the search, or write any comments you have about this event.		

Page. 1	Information-Event Card	Date: / / 2006
1. Information needed (what subject, in which format):		
2. How did you start looking for the needed information?		
<input type="checkbox"/> I. Asking a colleague/friend.		
If Yes, via: <input type="checkbox"/> A. Phone <input type="checkbox"/> B. E-mail <input type="checkbox"/> C. In person		
<input type="checkbox"/> II. Consulting printed resources:		
Resources: <input type="checkbox"/> A. Book <input type="checkbox"/> B. Journal <input type="checkbox"/> C. Others, name		
Location: <input type="checkbox"/> A. From library, <input type="checkbox"/> B. From your own collection		
<input type="checkbox"/> C. Other, name		
<input type="checkbox"/> III. Consulting electronic resources.		
<input type="checkbox"/> A. Electronic journal		
<input type="checkbox"/> B. Other, name		
<input type="checkbox"/> IV. General searching on the web		
<input type="checkbox"/> V. Searching a database, please name		
<input type="checkbox"/> VI. Other, please specify		
3. How did you proceed in the process of looking for the information (next steps, if any)?		

Continues overleaf →



## **Appendix 07: The invitation email for participation in the survey**

Dear...

My name is Hamid Jamali, and I'm researching scholarly communication and information-seeking behaviour of physicists and astronomers as part of a PhD thesis in the School of Library, Archive and Information Studies. I would be very grateful if you could take part in a questionnaire survey by filling the online questionnaire.

Pilot tests show that it will take approximately 6-10 minutes and you will be entitled to a £50 cash prize draw (just by entering your email address). The questionnaire is anonymous and your responses will be kept confidential.

The primary goal of this survey is simply to develop a better understanding of scholars' information needs and activities in order to improve information systems and services. As such, your answers are extremely important to the accuracy of my study. Please complete the survey within two weeks of receipt of this email.

To complete the questionnaire, please click on the following link or copy and paste it in your browser.

<http://www.soi.city.ac.uk/~ds137/b/q.htm>

Please feel free to contact me if you need to know more about this study. Or you can simply find out more on this page.

<http://www.homepages.ucl.ac.uk/~uczchrj/research.htm>

Let me take this opportunity to thank you for all your help in advance.

Kind Regards

Hamid

Ps. If you prefer to complete a print version of the questionnaire, just drop me a line/email and I will make sure a copy is delivered to your office.

## **Appendix 08: The reminder email for participation in the survey**

Dear...

Sorry to bother you again, but I thought I would send you a gentle reminder regarding my previous email (see below). If you could spare a few minutes to help me by filling out the questionnaire I would be very grateful. It's just that I haven't had a great response from the department, and the survey is crucial to my research.

If you have already filled it out, please ignore this email, and my apologies for bothering you again. This is the link:

<http://www.soi.city.ac.uk/~ds137/a/q.htm>

Many thanks again,

Hamid

## Appendix 09: The online questionnaire – staff version

1 What research group do you mainly belong to?

[List of research groups presented in a drop-down menu]

2 In your subfield, how important is rapid awareness of new papers?

- Not at all important
- A little important
- Somewhat important
- Quite important
- Absolutely critical
- I don't know

3 How dependent are you on each of these methods for keeping up-to-date with developments in your subfield?

(Scales: Very dependent, Quite dependent, Not very dependent, Not at all dependent)

- Browsing electronic journals
- Browsing print journals
- Browsing preprint archive
- Receiving journals' table of contents email alerts
- Receiving email alerts from preprint archives
- Receiving search email alerts (like the service of Web of Knowledge)
- Newsletters
- Departmental or groups' Seminars and meetings
- Conferences
- Word of mouth and colleagues
- Regular or semi-regular searching on a database or internet
- Other, please specify

4 Please rank the top three methods you depend on for keeping up-to-date.

- The most important method

- The second most important method
- The third most important method

5 How often do you use each of these methods for identifying research articles?

(Frequency: Daily, 2-3 times a week, About once or twice a month, Less than once a month, Never)

- Recommendation from friends
- Table of contents email alerts
- Browsing or searching journals' websites
- Following up references at the end of papers
- Searching in a general database such as Web of Knowledge
- Searching in a subject specific database such as ADS, Spires, Inspec
- Searching Google for words or authors (this doesn't include when you search Google to find a journal's website)
- Searching Google Scholar
- Other

6 By which of the above-mentioned methods do you identify the highest number of articles you read? Please write its number (the blue number) in this box.

### READING

7 Approximately how many published articles do you read each month? (By reading I mean going beyond abstract and reading some parts of the article at least).

8 Approximately how many preprints do you read each month?

9 Are you personally subscribed to any journal (Exclude those that you receive as a part of your society membership or you receive as editor)?

No Yes, please name ...

Please, think of the scholarly article you read most recently

10 How did you identify the last article you read?

- Through a colleague
- Through email alert
- Through browsing a journal's website

- Through references of another paper
  - Through a search on Google
  - Through a search on Google Scholar
  - Through a search in an abstract database (e.g. Web of Science, Spire)
  - I had read it before and was rereading it, so I already knew about it
  - Other, please specify
- 11 How old was that paper?
- A few weeks
  - A few months
  - 1 year
  - 2 years
  - 3 years
  - 4-5 years
  - 6-10 years
  - 11-15 years
  - More than 15 years

#### STATEMENTS

12 To what extent do you agree or disagree with the following statements about your information seeking activities?

(Scales: Strongly disagree, Disagree a little, Neither agree nor disagree, Agree a little, Strongly agree, I don't know)

- I tend to avoid using subject key words and phrases when searching databases to find articles because it brings up too many results.
- If an article is not available online, it's probably not worth the effort to obtain it.

#### PUBLICATIONS

13 In the past two years, how many articles in refereed journals have you published?

(Please do NOT include those that are in press)

14 Do you deposit most of your articles in an e-print archive such as arXiv.org?

Yes (move on to next question)    No (go to question 16)

15 When do you generally deposit your articles in e-print servers (such as arxiv.org)?

- Before I submit to journal
- When I submit to journal
- When accepted by a journal

16 What is the reason for not depositing your articles in e-print archives? (If you do deposit, move on to next question).

- Because I cannot be bothered
- Because I don't see any benefit in it
- Because it is not common or a tradition in my subfield
- Because the copyright of the journals I publish in doesn't allow it
- Other (specify)

### **YOUR SUBFIELD**

In this section I would like to know about your perception of your research area.

17 For your research, how often do you need to look for and use the results of research by people in other disciplines (e.g. chemistry etc.)?

Never Rarely Sometimes Often

18 What do you think of the journal literature of your subfield?

- It's very scattered in many journals and searchable through several databases
- It's reasonable, not very scattered and not very concentrated
- It is quite concentrated in a few journals and searchable through a few databases
- I don't know

### **PROBLEMS**

19 What practically causes you the most problems while seeking information and literature? (If any, please choose maximum two most important ones)

- The backfile and older issues of journals are not available online
- There are too many papers out there
- Obtaining papers from obscure journals
- Accessing electronic material from home
- Obtaining conference proceedings
- I don't have any problems
- Other, please specify

**DEMOGRAPHICS**

20 What is your gender?

Male      Female

21 Which best characterises your work (mainly)?

Physics: Theory    Experiment    A bit of both

Astronomy and astrophysics:

Observational data acquisition and analysis

Theory (including analytic or numerical modelling)

Instrumentation

22 What is your academic status? (including visitors, e.g. if you are a visitor lecturer please select Lecturer)

- Research assistant
- Research fellow
- Senior research fellow
- Lecturer
- Senior lecturer
- Reader
- Professor

23 How old are you?

34 and less    35-39    40-49    50-59    60 and more

24 How long have you worked at UCL?

1-2 years    3-5 years    6-10 years    More than 10 years

25 Which role have you taken during the last 4 years?

- Author of journal articles
- Referee of journal articles
- Editorial board member
- Guest editor
- None of the above

26 What percentage of your work time do you spend doing the following (the total should be 100%)?

- ..... % Research
- ..... % Teaching
- ..... % Writing (article, grant proposal etc.)
- ..... % Other

27 Finally, do you have any additional comments regarding your information seeking and communication activities (or concerning this study) you would like to mention?

28 If you would like to be entered into the £ 50 cash prize draw, enter your email address (don't put your name, just email) into the following box.



## Appendix 10: The online questionnaire - PhD students version

1 What research group do you mainly belong to?

[List of research groups presented in a drop-down menu]

2 In your subfield, how important is rapid awareness of new papers?

- Not at all important
- A little important
- Somewhat important
- Quite important
- Absolutely critical
- I don't know

3 How dependent are you on each of these methods for keeping up-to-date with developments in your subfield?

(Scales: Very dependent, Quite dependent, Not very dependent, Not at all dependent)

- Browsing electronic journals
- Browsing print journals
- Browsing preprint archive
- Receiving journals' table of contents email alerts
- Receiving email alerts from preprint archives
- Receiving search email alerts (like the service of Web of Knowledge)
- Newsletters
- Departmental or groups' Seminars and meetings
- Conferences
- Word of mouth and colleagues
- Regular or semi-regular searching on a database or internet
- Other, please specify

4 Please rank the top three methods you depend on for keeping up-to-date.

- The most important method

- The second most important method
- The third most important method

5 How often do you use each of these methods for identifying research articles?

(Frequency: Daily, 2-3 times a week, About once or twice a month, Less than once a month, Never)

- Recommendation from friends
- Table of contents email alerts
- Browsing or searching journals' websites
- Following up references at the end of papers
- Searching in a general database such as Web of Knowledge
- Searching in a subject specific database such as ADS, Spires, Inspec
- Searching Google for words or authors (this doesn't include when you search Google to find a journal's website)
- Searching Google Scholar
- Other

6 By which of the above-mentioned methods do you identify the highest number of articles you read? Please write its number (the blue number) in this box.

### READING

7 Approximately how many published articles do you read each month? (By reading I mean going beyond abstract and reading some parts of the article at least).

8 Approximately how many preprints do you read each month?

Please, think of the scholarly article you read most recently

9 How did you identify the last article you read?

- Through a colleague
- Through email alert
- Through browsing a journal's website
- Through references of another paper
- Through a search on Google
- Through a search on Google Scholar

- Through a search in an abstract database (e.g. Web of Science, Spire)
  - I had read it before and was rereading it, so I already knew about it
  - Other, please specify
- 10 How old was that paper?
- A few weeks
  - A few months
  - 1 year
  - 2 years
  - 3 years
  - 4-5 years
  - 6-10 years
  - 11-15 years
  - More than 15 years

#### STATEMENTS

11 To what extent do you agree or disagree with the following statements about your information seeking activities?

(Scales: Strongly disagree, Disagree a little, Neither agree nor disagree, Agree a little, Strongly agree, I don't know)

- I tend to avoid using subject key words and phrases when searching databases to find articles because it brings up too many results.
- If an article is not available online, it's probably not worth the effort to obtain it.

#### PUBLICATIONS

12 Do you (or would you) deposit most of your articles in an e-print archive such as arXiv.org?

Yes (move on to next question)    No (go to question 16)

- 13 When do you generally deposit your articles in e-print servers (such as arxiv.org)?
- Before I submit to journal
  - When I submit to journal
  - When accepted by a journal

14 What is the reason for not depositing your articles in e-print archives? (If you do deposit, move on to next question).

- Because I cannot be bothered
- Because I don't see any benefit in it
- Because it is not common or a tradition in my subfield
- Because the copyright of the journals I publish in doesn't allow it
- Other (specify)

### **YOUR SUBFIELD**

In this section I would like to know about your perception of your research area.

15 For your research, how often do you need to look for and use the results of research by people in other disciplines (e.g. chemistry etc.)?

Never Rarely Sometimes Often

16 What do you think of the journal literature of your subfield?

- It's very scattered in many journals and searchable through several databases
- It's reasonable, not very scattered and not very concentrated
- It is quite concentrated in a few journals and searchable through a few databases
- I don't know

### **PROBLEMS**

17 What practically causes you the most problems while seeking information and literature? (If any, please choose maximum two most important ones)

- The backfile and older issues of journals are not available online
- There are too many papers out there
- Obtaining papers from obscure journals
- Accessing electronic material from home
- Obtaining conference proceedings
- I don't have any problems
- Other, please specify

### **DEMOGRAPHICS**

18 What is your gender?

Male Female

19 Which year of your PhD are you in?

First   Second   Third   Fourth   Fifth+

20   Which best characterises your work (mainly)?

Physics: Theory   Experiment   A bit of both

Astronomy and astrophysics:

Observational data acquisition and analysis

Theory (including analytic or numerical modelling)

Instrumentation

21   Finally, do you have any additional comments regarding your information seeking and communication activities (or concerning this study) you would like to mention?

22   If you would like to be entered into the £ 50 cash prize draw, enter your email address (don't put your name, just email) into the following box.

## Appendix 11: Cluster analysis of questionnaire survey results

This appendix presents a cluster analysis of the questionnaire data. Cluster Analysis, also called data segmentation, relates to grouping or segmenting a collection of cases (individuals and so on) into subsets or 'clusters', such that those within each cluster are more closely related to one another than objects assigned to different clusters. The aim of the analysis was to find out whether participants from the same research group would fall in the same cluster indicating similarities in their information-seeking behaviour. To conduct the clusters SPSS software was used, as it was used for other statistical analyses presented in this thesis. Ward's hierarchical clustering method was used to create a horizontal dendrogram. In hierarchical clustering the resultant classification has an increasing number of nested classes. The cases were labelled by the research group they were a member in. Seven research groups were labelled using an alphabetical letter (see below). The variables used for clustering were all of the answers to questions 2-8 (see Appendix 9 for the questions).

A: Atmospheric Physics

B: High Energy Physics

C: Optical Science Laboratory

D: Condensed Matter and Material Physics

E: Astronomy and Astrophysics

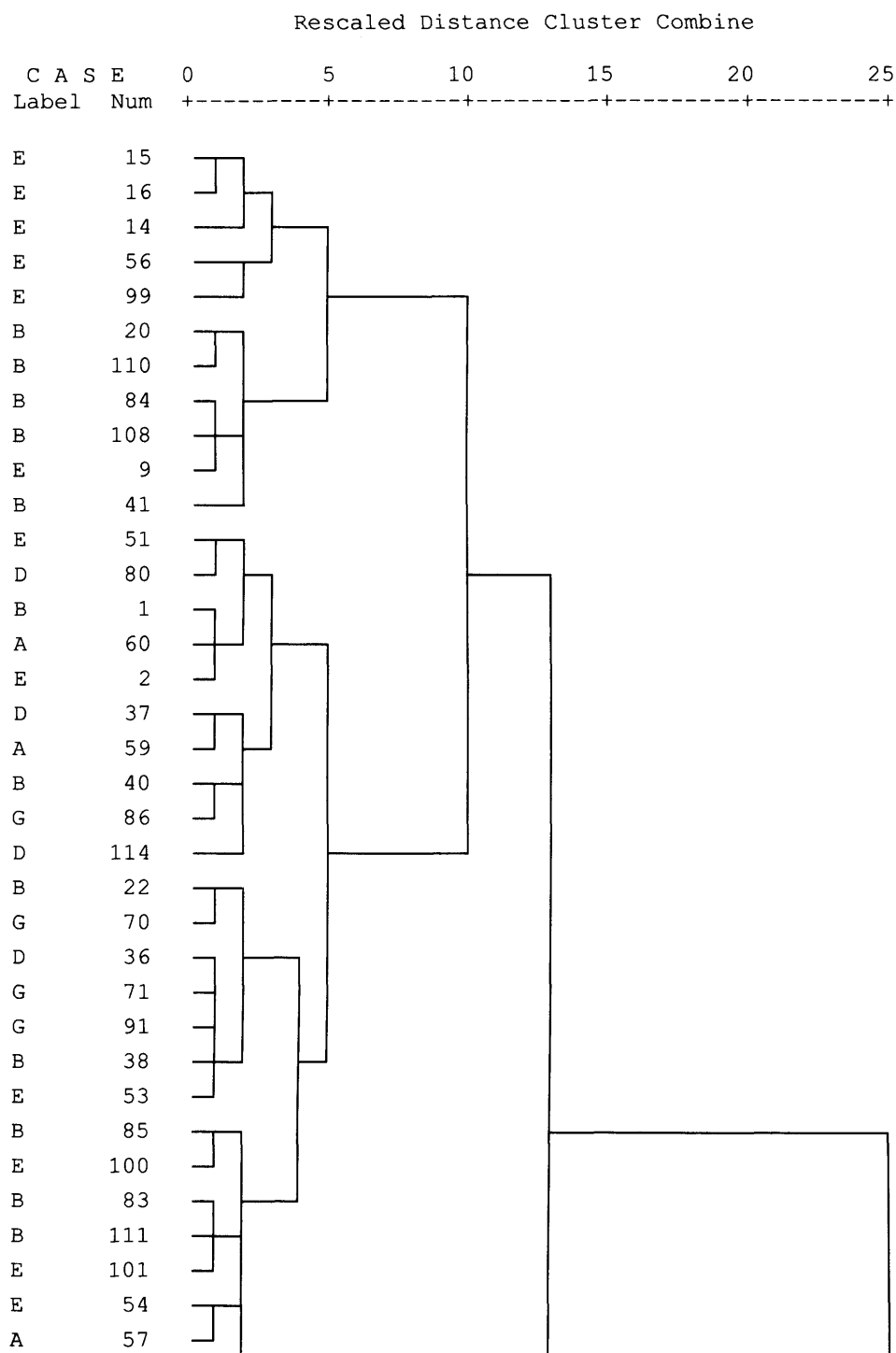
F: Theoretical Molecular Physics

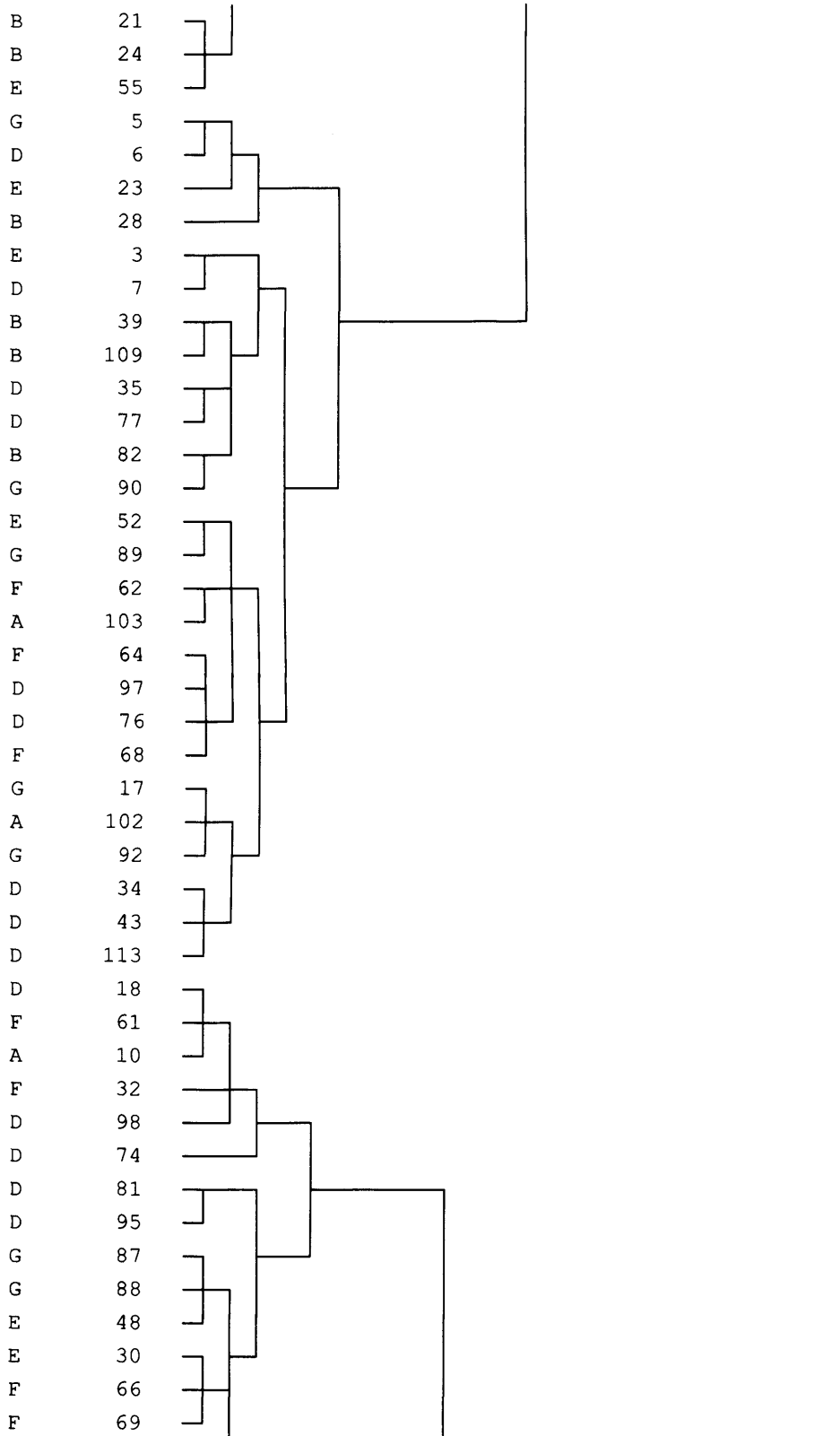
G: Atomic, Molecular, Optical and Positron Physics

As we can see in the dendrogram the distribution of cases does not indicate that people in the same research groups are very close considering the overall characteristics of their information-seeking behaviour. Respondents from different research groups are somewhat spread out in several clusters. This is not to say that there are no similarities in the information-seeking behaviour of respondents from the same research groups, because as it has been in the findings of the study there are similarities and even in the clusters one can identify people from the same groups in the same clusters. However, this analysis was done in order to help the researcher decide on the way of organising the results.

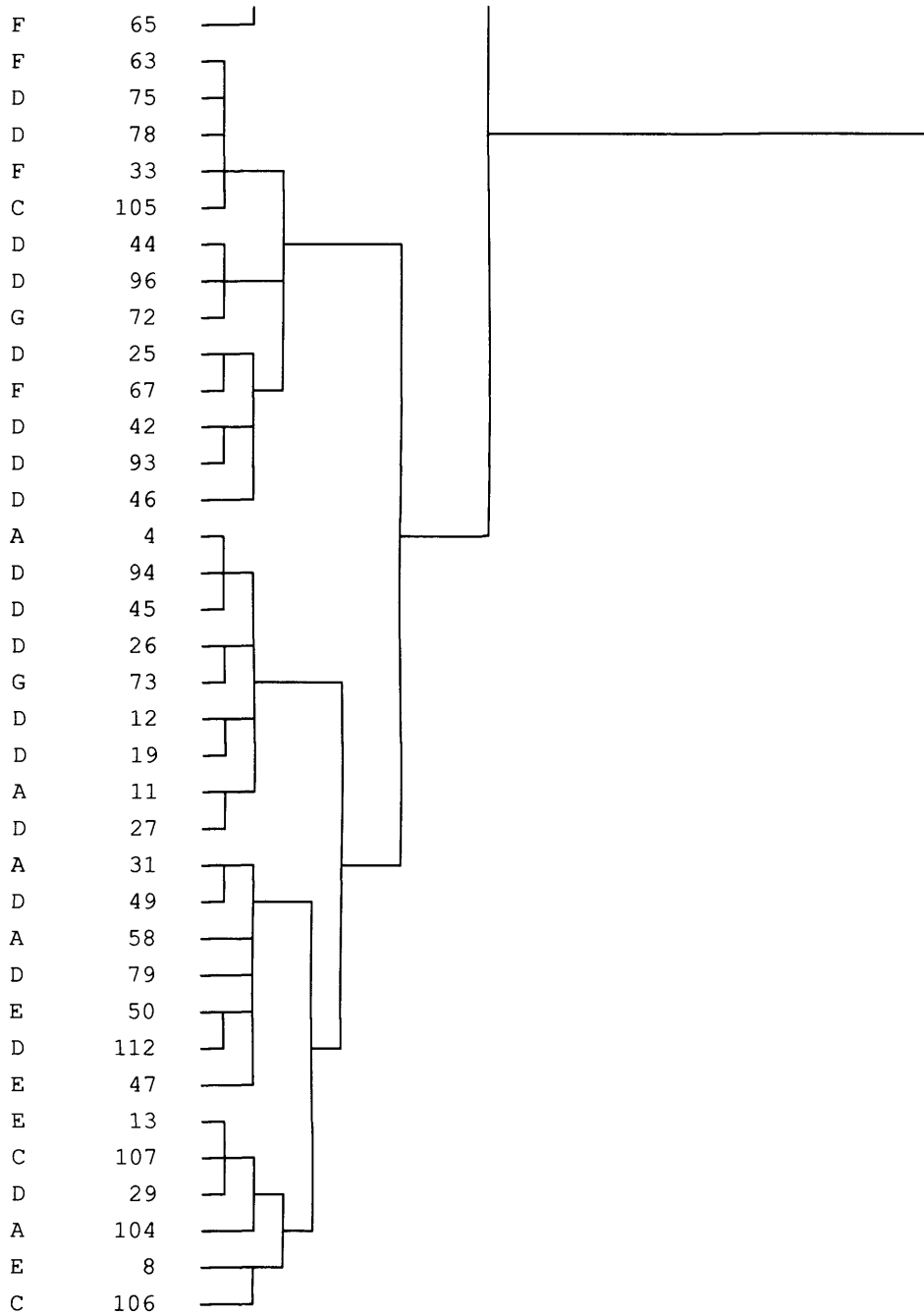
\* \* H I E R A R C H I C A L C L U S T E R A N A L Y S I S \* \*

Dendrogram using Ward Method









## Appendix 12: List of publications

As stated in the introduction chapter, this thesis fits into the CIBER's Virtual Scholar Research programme. The following list includes publications related to Virtual Scholar in which the researcher of the present thesis has contributed.

### 2008

Huntington, P., Nicholas, D. and Jamali, H. R. (2008). Web Robot detection in the scholarly information environment. *Journal of Information Science*, **34** (accepted, in press)

Nicholas, D., Huntington, P. and Jamali, H. R. (2008). User diversity: as demonstrated by deep log analysis. *Electronic Library*. **26** (1), 21-38.

Huntington, P., Nicholas, D. and Jamali, H. R. (2008). Web site usage metrics: a re-assessment of session data. *Information Processing and Management*, **44** (1), 358-372.

### 2007

Rowlands, I., Nicholas, D., Jamali, H. R. and Huntington, P. (2007). What do faculty and students really think about e-books? *Aslib Proceedings*, **59** (6), 489-511.

Huntington, P., Nicholas, D. and Jamali, H. R. (2007). Employing log metrics to evaluate search behaviour and success: case study BBC search engine. *Journal of Information Science*, **33** (5), 584-597.

Huntington, P., Nicholas, D. and Jamali, H. R. (2007). Site navigation and its impact on content viewed by the virtual scholar: a deep log analysis. *Journal of Information Science*, **33** (5), 598-610.

Nicholas, D., Huntington, P. and Jamali, H. R. (2007). Diversity in the information-seeking behaviour of the virtual scholar: institutional comparisons. *Journal of Academic Librarianship*, **33** (6), 629-638.

Nicholas, D., Huntington, P. and Jamali, H. R. (2007). Open access in context: a user study. *Journal of Documentation*, **63** (6), 853-878.

Nicholas, D., Huntington, P. and Jamali, H. R. (2007). The use, users and role of abstracts in the digital scholarly environment. *Journal of Academic Librarianship*, **33** (4), 446-453.

Nicholas, D., Huntington, P., Jamali, H. R. and Dobrowolski, T. (2007). Characterising and evaluating information seeking behaviour in a digital environment: spotlight on the

'bouncer'. *Information Processing and Management*, **43** (4), 1085-1102.

Huntington, P., Nicholas, D., Jamali, H. R. and Russell, C. (2007). Health information for the consumer: NHS vs the BBC. *Aslib Proceedings*, **59** (1), 46-67.

Nicholas, D., Huntington, P. and Jamali, H. R. (2007). The impact of open access publishing (and other access initiatives) on use and users of digital scholarly journals. *Learned Publishing*, **20** (1), 11-15.

## 2006

Huntington, P., Nicholas, D., Jamali, H. R. and Watkinson, A. (2006). Obtaining subject data from log files using deep log analysis: case study OhioLINK. *Journal of Information Science*, **32** (4), 299-308.

Huntington, P., Nicholas, D., Jamali, H. R. and Tenopir, C. (2006). Article decay in the digital environment: a usage analysis by date of publication employing deep log methods. *Journal of the American Society for Information Science and Technology*, **57** (13), 1840-1851.

Jamali, H. R. and Nicholas, D. (2005). "Communication and information-seeking behaviour of PhD students in physics and astronomy". In *Proceedings of the American Society for Information Science and Technology*, **43** (1), 106-124.

Nicholas, D., Huntington, P., Jamali, H. R. and Tenopir, C. (2006). What deep log analysis tells us about the impact of big deal, Case study OhioLink. *Journal of Documentation*, **62** (4), 482-508.

Williams, P., Jamali, H. R. and Nicholas, D. (2006). Using ICT with people with special education needs: what the literature tells us. *Aslib Proceedings*, **58** (4), 330-345.

Nicholas, D., Jamali, H. R. and Rowlands, I. (2006). On the tips of their tongues: authors and their views on scholarly publishing. *Learned Publishing*, **19** (3), 193-203.

Nicholas, D., Huntington, P., Jamali, H. R. and Watkinson, A. (2006) "The information seeking behaviour of the users of digital scholarly journals", *Information Processing and Management*, **42** (5), 1345-1365.

Nicholas, D., Huntington, P., Jamali, H. R. and Tenopir, C. (2006). Finding information in (very large) digital libraries: a deep log approach to determining differences in use according to method of access. *Journal of Academic Librarianship*, **32** (2), 119-126.

**2005**

Jamali, H. R., Nicholas, D. and Huntington, P. (2005). Use and users of scholarly e-journals: a review of log analysis studies. *Aslib Proceedings*, **57** (6), 554-571.

Nicholas, D., Huntington, P., Russell, B., Watkinson, A., Jamali, H. R. and Tenopir, C. (2005). The Big Deal – ten years on. *Learned Publishing*, **18** (4), 251-157.

Nicholas, D., Huntington, P., Dobrowolski, T., Rowlands, I., Jamali M., H. R., Polydoratos, P. (2005). Revisiting ‘obsolescence’ and journal article ‘decay’ through usage data: an analysis of digital journal use by year of publication. *Information Processing and Management*, **41** (6), 1441-1461.

Nicholas, D., Jamali, H. R., Huntington, P. and Rowlands, I. (2005). In their very own words: authors and scholarly journal publishing. *Learned Publishing*, **18** (3), 212-220.

THE END