The Possible Effects of a Bibliometrically-Based Evaluation System on Research Behaviour: The Case of the University of Iceland

Eva Dögg Diego Porkelsdóttir
Acknowledgements

I want to use the opportunity to acknowledge the help of all the people, who have directly or indirectly, contributed to the production of this dissertation. Foremost, I would like to express my gratitude for the invaluable help and guidance of my supervisor, Birger Larsen, at the Det Informations-videnskabelige Akademi in Copenhagen. I also want to thank Prof. Peter Willett at the University of Sheffield for his priceless help with the initial construction of the literature review. Furthermore, I would like to give special thanks to the staff members of the Division of Analysis, Evaluation and Indicators at the Icelandic Centre for Research for their assistance and advice.
Abstract

The academic world of science and research has had a boost in the demand and application of measuring activities intended to translate science, in its diverse and multiple forms, into numbers and sometimes monetary amounts. One of these measuring tools are bibliometrically-based evaluation systems, which focus on, but are not limited to, publication and citation analyses. An example of such a system is the current Performance based evaluation system of public universities, used at the University of Iceland.

The purpose of the dissertation was to examine how such a system could potentially affect the research behaviour of staff members at the university. Consequently, the first objective was to examine bibliometric numbers before and after the implementation of the system in a comparative manner. This part of the analysis included a comparison to six other Nordic universities, which were not applying the same system, in order to see if the same or similar trends occurred. The second objective was to examine the views of the personnel at the UoI, in order to receive an idea of how they experienced or perceived the evaluation system. These research strategies yielded comparable information on publications, citations, collaboration, system equality and research behaviour.

The study did not reveal any significant bibliometric changes in publications, citations or collaboration, which could be directly linked to the implementation of the evaluation system. In addition, a publication format analysis did not reveal an inherent inequality within the system, although, it must be noted that inequality could stem from other systematical aspects. The questionnaire results did to some degree reflect a different reality for the university’s staff members, as the they perceived the system to stimulate research, but simultaneously be inherently unequal. Staff did, however, agree with the collaboration analysis, as it did not see the system as a diminishing factor in collaboration. Lastly, around 40% of respondents would like citations to be more actively used within the system.
# Table of Contents

**Introduction** .................................................................................................................................................. 1

**Chapter 1 – Methodology** ................................................................................................................................. 4
  1.1 Theoretical Construction and the Framework of Research ......................................................................... 4
  1.2 Research Design and Strategies ................................................................................................................. 11
  1.3 Summary .................................................................................................................................................... 18

**Chapter 2 - Literature Review** .......................................................................................................................... 20
  2.1 Bibliometric Indicators – Appliance and Possible Issues ........................................................................... 20
    2.1.1 Publication and Citation Analysis ....................................................................................................... 20
    2.1.2 The Theoretical Aspect ....................................................................................................................... 21
    2.1.3 The Technological Aspect .................................................................................................................. 22
    2.1.4 The Methodological Aspect and Possible Pitfalls in Appliance ......................................................... 24
  2.3 Bibliometrically-Based Research Evaluation Systems ................................................................................ 27
    2.3.1 Using Publication Counts and Determining Impact ........................................................................... 28
    2.3.2 Allocating Credit amongst Contributors – Possible Effects on Collaboration .................................. 31
    2.3.2 Applying the Models to All Disciplines – Questions of Equality ...................................................... 31
  2.4 Research at the University of Iceland and the Icelandic Performance Based Evaluation System .................... 34
    2.4.1 The Icelandic Performance Based Evaluation System ...................................................................... 37
  2.5 Summary .................................................................................................................................................... 42

**Chapter 3 – Bibliometric Analysis Results** ........................................................................................................ 43
  3.1 Publication Analysis ....................................................................................................................................... 43
  3.2 Collaboration ............................................................................................................................................... 49
  3.3 Citation Analysis ......................................................................................................................................... 52
  3.4 Research Points and Publication Format ................................................................................................... 56
  3.6 Summary .................................................................................................................................................... 62

**Chapter 4 – Questionnaire Results** .................................................................................................................. 63
  4.1 Population and Response Rates .................................................................................................................. 64
  4.2 Demographic Questions .............................................................................................................................. 65
  4.3 Satisfaction with the Performance-Based Evaluation System ..................................................................... 70
  4.4 Publication, Quality and Citations .............................................................................................................. 71
  4.5 Collaboration ............................................................................................................................................. 76
  4.6 Equality ...................................................................................................................................................... 80
  4.7 Research Behaviour .................................................................................................................................... 85
4.8 A New System .................................................................................................................. 87
4.9 Summary ......................................................................................................................... 90

Chapter 5 – Discussion ........................................................................................................... 92
  5.1 Publication Analysis .......................................................................................................... 92
  5.2 Collaboration .................................................................................................................... 96
  5.3 Citation Analysis .............................................................................................................. 97
  5.4 Research points and Equality .......................................................................................... 100
  5.5 Impact on Research Behaviour ....................................................................................... 102
  5.6 The Overall Satisfaction and a New System? ................................................................. 105
  5.7 The Theoretical Impact Model ....................................................................................... 105
  5.8 Final Observations and Recommendations ................................................................... 108
  5.9 Methodological Reflections ........................................................................................... 110

Conclusion ............................................................................................................................. 111

Bibliography ............................................................................................................................ 114

I. Appendix – Questionnaire Layout ...................................................................................... 119
II. Appendix – Additional Visual Survey Results ................................................................. 127
III. Appendix – Evaluation System for Public Universities .................................................. 131
Introduction

Within the academic world of science and research the interest and demand for measuring academic activity and impact has grown extensively, which can, for example, be viewed by looking at the popularity of university ranking and research evaluation. This development has brought about the creation of evaluation systems, whose goal it is to encompass science in its diverse and multiple forms for the purpose of making it measurable, both in numbers and sometimes monetary amounts.

Consequently, the aforementioned development has also seen the rise of bibliometric analyses being used for structuring such evaluation systems, as well as monitoring them. As a result, the area of library and information science has been enlarged even further and provided with the opportunity to intertwine its research with those of evaluation research, policy making and even business.

An example of such an evaluation system is the Performance-based Evaluation System (PBES) current being used at the University of Iceland (UoI). The system was introduced in 1998 and has since been widely debated within the academic community and can be described as having a rather controversial status. Its role is to evaluate the research partaken within the institution, as well as to encourage further research and personally award those deemed successful in this role by the system.

Even though the system has been discussed within the academic community, no coherent analysis or investigation has been done to assess or judge the system’s merit or function. At least no such evaluation has been published or made readily available, which presents a certain knowledge gap the author of this study was fully made aware of during the study’s conduction. Nevertheless, a single report on the scholarly work at the UoI was published in 2005 (Sigfúsdóttir, Ásgeirsdóttir, Macdonald & Feller, 2005), which
incorporated a discussion on the evaluation system, as well as bibliometric studies on a national level conducted by Sigfúsdóttir & Þórlindsson (2000) and the Icelandic Centre for Research (2010) are available. Apart from these publications, the field of bibliometric research in Iceland remains quite barren.

Monitoring the evaluation system more closely would have provided bibliometricians and policy makers with a wealth of information, especially useful when developing the system further. The information could also have been useful for those structuring a new system in other institutions or countries, due to the university’s experience of an evaluation system based on a personal award approach.

The purpose of this study will, therefore, be to commence on filling up this knowledge gap with a bibliometric, as well as other method driven, analysis of the PBES at the UoI. The aim will be to examine how or if the implementation of the system has brought about changes in research behaviour of those affected by it.

Consequently, the first objective will be to investigate the changes in bibliometric numbers before and after the implementation of the system in a comparative manner. This will be done by observing other Nordic universities not making use of the same system, and see if the same or similar trends occur. The second objective will be to examine the views of staff members at the UoI in order to receive an inkling of their experiences and to detect if the personnel’s views in some manner reflect those proposed by the bibliometric analysis.

The structure of the study will be as follows:

Chapter 1 outlines the theoretical, as well as methodological, approaches of the study. The dissertation draws upon the theory-driven programme theory as described by Stuart I. Donaldson (2007), which will provide a framework for the analysis. The triangulation of methods is also introduced, that is, the use of interviews, a survey and bibliometric methods.
Chapter 2 explores various literature on bibliometric indicators and their use in research evaluation, as well a discussion and introduction to various bibliometrically-based evaluation systems, including the Icelandic one. Chapter 3 outlines the results of the bibliometric analysis and the survey results are presented in Chapter 4. Chapter 5 will include a discussion of the aforementioned results, as well as their relation to the theoretical model and the literature review. The dissertation will then end with a conclusion, which will bring together the final results, as well as recommendations for further research.
Chapter 1 – Methodology

As the purpose of this study is to investigate both perceptions, as well as quantitative programme outcome, it is necessary to apply different methods of research to capture the desirable data. The objective of this chapter will, therefore, be to outline the theoretical construction of the investigative approach and the research framework, along with the research strategies and their design.

1.1 Theoretical Construction and the Framework of Research

As was mentioned in the introductory chapter, the aim of this study is to research and evaluate a specific “real world” programme, which presently is being used by a large tertiary educational institution. It is, therefore, clear that the study is more at home with the applied sciences than the basic sciences, although the difference between the two is more of a practical, rather than methodological, nature:

_Evaluation research is one type of applied research. Its purpose is to answer practical real-world questions about the effects of some policy or program…evaluation research has as its primary goal the description and evaluation of some existing social policy or program_ (Hoyle, Harris and Judd, 2002, p. 338).

As the nature of this study is in fact evaluative and perfectly suited with the concept of evaluation research described above, the question is not whether to classify it as evaluation research, but rather to decide which evaluation approach to select, that is, which one would serve this specific study best.

After having reviewed the various literature on evaluation research (Donaldson, 2007; Hoyle, Harris and Judd, 2002; Rossi, Freeman and Lipsey, 1999; Hall & Hall, 2004; Chen, 1990; Patton, 1990), it was established that the choice of evaluation approaches would be between either a conventional methodological approach or a theory based approach. In
general, evaluation research is method driven, that is, grounded in the methodology or tools used for each evaluation, resulting in a variety of approaches (see overview e.g. in Donaldson, 2007 and Hall & Hall, 2004). It was, however, in the 1990s that evaluators started to develop theories of evaluation; constructing a theoretical base for evaluation research, and criticising the lack of theory in the method driven evaluation models. Those new evaluation theories included constructive and realistic approaches, as well as theory-driven evaluation (Hall & Hall, 2004).

When deciding which approach to use for this study, the determining factors were established as being first and foremost the limited scope of the evaluation, and the fact that the programme under evaluation is presently being implemented in a real life settings of a semi-large educational institution. The fact that the work would be undertaken within the uncontrollable, and perhaps unstable, settings of daily life, made the choice of a strong theoretical framework more feasible. As a result, the choice became between the various theory-based approaches.

Determining the right approach also needed to take into account the fact that the study would be on a small scale. What is meant by small scale is that due to the short time span of the study (maximum five months), the evaluation research will only be able to touch upon and incorporate a part of the whole programme. When evaluating how the different theory-based evaluation techniques would fit this narrow frame, the three major approaches mentioned above were reviewed.

The constructive theory evaluation introduces the idea of a relative approach, where the programme under evaluation is perceived as being made up of different perspectives or constructions of all stakeholders (Hall & Hall, 2004). The evaluation, therefore, focuses on collecting all the possible constructions as described by Hall & Hall: “It is the task of the
evaluator to create a discourse with the many different stakeholders in a program…in the hope of reducing the number of constructions to one agreed pattern.” (2004, p. 53) As interesting as this approach would be for this study, the time scope would simply not allow for an elaborate discourse between all different stakeholders, who, theoretically, would not only be the university’s staff and administration, but also the Icelandic government and society.

Realistic evaluation, yet another possible approach, is a theory-based evaluation, which sees the programme as an individual entity, both present and researchable, unlike the beliefs of the constructivists above (Hall & Hall, 2004). The focus of the realistic approach is on causal patterns within the programme, which are called ‘mechanisms’. These mechanisms are seen as causing patterns or some sort of regularities, which, in combination with a particular context, will explain the outcome of the programme (Hall & Hall, 2004). As a result, receiving a coherent picture of the programme under evaluation the approach calls for an extensive evaluation, where the different mechanisms and different contexts are investigated (Hall & Hall, 2004). It is, therefore, greatly beyond the scope of this particular study to attempt a realistic evaluation in the spirit of realistic approach.

The last example of a theory-related evaluation is the one of programme theory-driven evaluation. This particular type of evaluation models offers both a flexible, as well as firm, approach to evaluations, without any possible time constraints. The flexibility of the approach is evident from its methodologically neutral position, as well as inclusive definition:

Program Theory-Driven Evaluation Science is the systematic use of substantive knowledge about the phenomena under investigation and scientific methods to improve, to produce knowledge and feedback about, and to determine the merit, worth and significance of evaluands such as social, educational, health, community and organizational programs (Donaldson, 2007, p. 9).
In this case, the Program Theory-Driven Evaluation would be applied to an educational/organisation programme, in order to produce knowledge and attempt to determine the effects of the system on the research behaviour of academics.

Contrastingly, the encompassing theory-driven framework provides a firm hold of the study (Donaldson, 2007). The research frame is constructed using a three step model, where the study is laid out in an organised manner:

1. *Developing program impact theory.*
2. *Formulating and prioritizing evaluation questions.*

These questions will guide the evaluator through the most important aspects of the study, whereas the flexibility mentioned above will make sure that whichever methodological approach suitable for the task, as well as whichever programme evaluated, becomes a valid choice. The combination of these two contrasting, but also counterbalancing elements of the theory-driven approach, is, therefore, the reason for the choice of Program Theory-Driven Evaluation for this study.

The first step of constructing the research framework, will, therefore, be to develop a programme impact theory. Firstly, it must be noted that a programme impact theory does not strive to explain the inner mechanisms of the programme, but is focused on the outcomes:

*Program impact theory, on the other hand, describes the cause-and-effect sequences that link the program services and activities to proximal and distal outcomes. A typical program impact theory illustrates how a program is expected to result in important desired outcomes, solve the problems of interest, or meet the needs of its target population* (Donaldson, 2007).
It must, however, be noted that in this particular study the discussion reaches beyond the mere outcomes, as the reasons for the wanted or unwanted outcomes are also investigated. It must, therefore, be concluded that the discussion in subsequent chapters stretches into the realm of programme process theory, that is, the theory of what happens within the system or the “black box” (Donaldson, 2007). Although, this discussion will always be in relation to the programme impact theory developed below.

Programme impact theory is most often pictorially described using a variable oriented approach (Donaldson, 2007), where the variables and their relationships are determined for further assessment. A simple illustration of a programme’s impact is viewable left. Here the programme is represented by the square furthest left and the ultimate outcome the mediator, which represents a variable “affected by the program, which in turns affects the outcome of interest” (Donaldson, 2001). Lastly, there is the moderator variable, which can affect relationships between the aforementioned variables, either by strengthening or weakening them. The moderator can affect the relationship between programme and mediator, as well as the mediator and outcome (Donaldson, 2001).
Applying this basic model to the Icelandic PBES would result in the following pictorial representation of a programme impact theory.

Here the system is suppose to motivate staff (mediator/proximal outcome) to publish more publications within prestigious publications channels (mediator/proximal outcome) as means of furthering arts and sciences (distal outcome). The other distal outcome of the system would be higher salaries, as well as higher status for the staff members. This simple model does certainly not capture all the possible affecting agents, but does give a concise and clear idea of what the programme’s aims are and how they are to be achieved.

The questions this impact theory poses are quite a few. Looking at the model it is evident that it tries to change the behaviour of researchers in a specific way, but what if the model is, due to its infrastructure and organisation, somehow altering the connections between the variables, perhaps even to the degree of the programme being non-functional in terms of its aims?
In order to try to answer this question, the aims of the study are threefold, that is, to examine the publications and citations as proximal outcomes (the blue ring), the perceived inequality, the effect of the system on research collaboration and other perceived ideas of the system’s function as a moderator (the yellow explosion), and finally the system itself in order to detect how the results from the abovementioned examinations can be explained in light of the system overall infrastructure and organisation.

For the purpose of capturing the different variables presented within the programme impact theory, varying research approaches were determined to be the most feasible option. The use of methodological triangulation was, therefore, chosen, resulting in the methods of interviewing, survey implementation and bibliometric analysis being used. As a result, the data gained from these three methods were both qualitative (interviews and parts of survey) as well as quantitative (survey and bibliometric numbers). The reason for choosing to address the PBES from different angles via various approaches is rooted in the notion that such a
 triangulation be aiding in the thorough depiction of the programme in question, as well as contribute to a more convincing and precise conclusion of this study (Yin, 1994).

1.2 Research Design and Strategies

As one of the aims of this study is to investigate stakeholder perceptions of the PBES, it is essential to apply appropriate research methods in order to approach the issue in the most fruitful manner. Reaching the population, in this instance the academic staff members of the UoI, and their perceptions, could be done in a number of ways. Those include either paper and pencil/internet questionnaires or face-to-face/telephone interviews.

Firstly, whether to use a qualitative interview approach or a more quantitative oriented survey approach to reach the desired population needed to be determined. The pros of using the face-to-face interview approach are the high response rate, the highest any mode of measurement can offer, and the ability to clarify questions or probe for further information, as well as control the context in which the questioning takes place. Furthermore, this method is also useful when trying to reach special populations, such as those not approachable by phone or e-mail due to socioeconomic reasons (Hoyle, Harris and Judd, 2002). Apart from this last aspect, as well as the ability to offer visual aid, the telephone interview has the same advantages as the face-to-face interview (Hoyle, Harris and Judd, 2002).

Despite the aforementioned merits the interview method introduces a set of possible disadvantages. Conducting a set of interviews can be both costly and time consuming (Hoyle, Harris and Judd, 2002). Additionally, there is always the inherent danger of social desirability, whereby respondents answer questions in a manner they believe to be socially acceptable (Trochim, 2006), and interviewer distortion, where the interviewer’s expectations or characteristics could potentially effect responses (Hoyle, Harris and Judd, 2002).
In this particular study the most limiting factor is the one pertaining to time. As one of the objectives is to collect perceptions of all available academic staff members, the idea of conducting roughly 700 interviews is not only daunting, but also impossible within the time scope of this dissertation. Another approach would be to construct a random quota sample from the stakeholder population, either in the form of individual interviews or even focus groups. This method would, none the less, be very time consuming as finding a representative group, that is, a group which would not only represent different university departments, but also gender, age and status within the institution, and then conduct interviews with all applicable members would simply take too much time.

Having established the time consuming nature of the interview approach, and thereby excluded it as a method of measurement for the stakeholder part in this dissertation, it was time to take a closer look at the survey method. In terms of methodology the difference between paper-based and internet-based surveys is minimal (Dillman, Smyth and Christian, 2009). It is, however, in the implementation that the differences become apparent.

Structuring a small scale internet-based survey is relatively easy, as various internet survey providers offer a variety of user friendly interfaces. Another advantage, and one of the most pertinent to this specific investigation, is that responses in internet-based surveys are more quickly turned in, much quicker than those of paper-based surveys (Sue & Ritter, 2007). In addition, survey services can be free or reasonably charged, and the distribution of the survey link easy via e-mail or websites. This is, however, not the case with paper-based surveys, as printing and postage costs can be considerable.

After having established the clear advantage of using an internet-based survey for this particular element of the research, it must however, be noted that no method is without its flaws. In general a web-based survey can introduce the bias of underrepresentation; excluding
participants who are not literate, in the conventional meaning of the word, as well as people who are not computer literate. Additionally, the survey requires people to be able to access a computer which can be connected to the internet (Hoyle, Harris and Judd, 2002). Nevertheless, it is safe to assume that this form of socioeconomic bias is not a pertinent factor in this particular study, as the population consists of highly educated individuals, who are dependent upon both computers and the internet for their work.

Lastly, other drawbacks of web-based, as well as paper-based, surveys must be noted. Firstly, there is always the possibility that someone else than the determined respondent answers the questionnaire (Hoyle, Harris and Judd, 2002). Secondly, there is the problem of low and incalculable response rates. As noted by Hoyle, Harris and Judd (2002), the response rate in web-based surveys is often poor, and, in some instances, determining how many actually answered the survey can be difficult.

Again, some of these drawbacks are not a determining factor in this study. Firstly, the invitations to participate in the survey were sent to specific e-mail addresses of a predetermined population, thereby minimising the possibility of unwanted respondents, and making response rate calculations effortless. The disadvantage of a low response rate is however, a difficult problem to solve. It must, therefore, be noted that this particular con is the prize which must be paid for the important advantage of quicker implementation within a short time frame.

Having determined the internet-based survey as the most convenient method of measurement, one of the stakeholders, in this case the academic staff of the UoI, was presented with a web-based questionnaire. The questions therein were designed as to reflect the views and experiences of the staff members regarding the desired and undesired outcomes of the Icelandic PBES.
The survey was conducted online from the 8th to the 25th of October using an online survey service from Questionpro. The survey was sent out via e-mail to all academic staff at the UoI, resulting in 782 possible respondents. The population was determined by collecting staff e-mails of the university’s website, where only academic staff with research obligations were chosen. Some of the e-mail addresses had, however, not been updated, or were simply not in use any more, which decreased the number of respondents, ending with a final population number of 776.

The questions in the survey were formed as to address the desired outcome of the PBES; increased production and impact of research, as well as possible unintended outcomes in relation to collaboration and possible discrimination. In addition, the survey was laid out in two languages, that is, Icelandic and English.

The first part of the survey consisted of five demographic questions regarding the participants’ status within the university; professor, lector, docent, adjunct or other, as staff members could possibly be classified as “specialists” or simply “researchers”; to which academic department or research institution they belonged; the percentage of their research obligation; and finally their gender and age.

Next, respondents were presented with Likert-scale questions regarding the respondents’ perceptions of the PBES, where they were asked how strongly they identified themselves with various statements regarding the effects of the PBES. In between respondents were asked about their opinion in a qualitative open-ended manner, as to especially elaborate further on various issues. The survey then ended with an open-ended question, prompting the respondents for further comments on the PBES.

The other part of the study centred on bibliometric analyses, with the focus point being placed on the PBES. Evaluating the PBES in itself was one of the goals, as well as
incorporating a comparative view with other universities. This was thought of as a tool to identify any specific changes, should they be present, which might be linked in an indirect manner to the implementation of the PBES. The universities used in comparison with the UoI were the Universities of Bergen, Tromsø, Lund, Uppsala and Copenhagen.

The reason why these universities were chosen for comparison is grounded in the UoI’s strategy for 2006 to 2011, where a wish for a comparison with these specific institutions (along with Boston University and the University of Aberdeen) is expressed (The University of Iceland, 2006a). The University states the reason for choosing these institutions to be the following:

*These comprise, on the one hand, several universities which the University of Iceland was compared to in the Government Audit Department's audit of 2004, which are comparable with the University in terms of operational form, size or range of study programmes. On the other hand, several universities have been chosen in the Nordic countries and in the USA, which are ranked among the 100 best in the world* (The University of Iceland [UoI], 2006a).

Why Boston University and the University of Aberdeen were left out of this study is due to two reasons. Firstly, due to the limited scope of the study not all university could be included. Secondly, when trying to capture possible effects of the PBES, isolating the system as the main contributing variable is merely impossible. Choosing universities in other Nordic countries, which share similar values and have a similar societal infrastructures, would perhaps eliminate some unwanted cultural effects on research performance. This is not to say that because of this selection a cultural, or any other nationally specific effect, is not present, but when determining the comparativeness of the institutions this societal likeness was seen as a possible restraint.
The year span used for the analyses was generally from 1994 to 2002. The reason for selecting this time period over any other, was for comparative purposes, that is, as it represents the four years before the implementation of the PBES, the year when PBES was implemented - but no experience or effect on salaries was still present, and the four years after the implementation. There were, however, some instances where data was only available for other year spans than those mentioned above. In those situations the available data was used, but note made of a change in time range.

The platform used for the analysis was Thompson Reuters’ Web of Science (WoS) – all databases. Wherefore this database was used instead of others, such as Scopus or Google Scholar, was twofold. Firstly, the practical issue of access only allowed for analyses within either WoS or Google Scholar. Secondly, when viewing the UoI’s PBES a great emphasis has been placed on ISI publications (now Thompson Reuters). As a result, the WoS was found to be the more appropriate choice of the two. It must, however, be noted that a further analysis for the arts and humanities, as well as the social sciences, using GS might be a feasible option due to the limited coverage of such material within the WoS (Moed, 2005).

The first part of the bibliometric study included a publication analysis of all six universities. All academic publications were collected for each year during the time span of the nine years for each institution. Furthermore, those publications were divided into subject categories, as means of creating a subject oriented profile of each institution, as well as offering the possibilities of further analyses. Not all 250 subject categories of the WoS were used, but only those in which the UoI had any publications, resulting in 195 categories. The reason for not using all subject categories in WoS is simply related to the limited time and scope of the study. In addition, for the purpose of an easier comparison, those categories were then merged together into 14 head groups based on the Noria-Net categories (NordForsk, 2010).
When those numbers had been retrieved it soon became apparent that the institutions being compared were of different sizes, at least in terms of publications. The best way to normalise the publication count would have been to obtain number of active researchers at each university for each year. Finding those numbers for all institutions and all years would, however, have been too difficult and time consuming for this project. Furthermore, population numbers for each nation would have been easier to come by, but would, however, have provided a skewed picture due to the fact that the universities differ in size, even those within the same country.

As a result, the normalisation technique applied was the one of index calculations, which provided a picture of the growth of publications within each institution. This meant that all universities had the same starting point of 100 in 1994 and from there growth was calculated upwards or downwards towards 2002.

A second component of the publication investigation included a collaboration analysis. This part of the study sought to determine the degree of collaboration of all six universities by counting the number of articles containing a foreign collaboration within the total publication count. This was done by ranking publications by country, subtracting those purely national, and thereby obtaining the number of articles containing collaboration.

Secondly, a citation analysis was conducted using the publications determined by the aforementioned publication analysis. Firstly, the citation impact of all universities for all nine years was determined. The citation window ranged from the year of publication to the time of the study’s execution, that is, December 2010. This provided each article with roughly 8 to 16 years to collect citations.

In addition, the citation numbers were normalised, as it has been shown that citation behaviour differs and might not be the same for all fields of research (Moed, Burger,
Frankfort & van Raan, 1985). The normalisation was obtained by calculating the Weighted Field Crown Indicator (WFCI) (Moed, Debruin & van Leeuwen, 1995) for each university per year of the nine year time period. This was done by obtaining the world citation impact of all the 195 ISI categories, using the National Science Indicator, and then compiling those numbers into the 14 Noria-Net categories. This way the actual citation impact of each Noria-Net category was made comparable to an expected impact, one determined by a world citation impact baseline. Hence, the WFCI could be calculated for each year, resulting in a compatible WFCI score for each institution for comparative means.

The last aspect of the bibliometric analysis focused on the research points granted by the UoI to departments and academic schools within the institution. These numbers were obtained from the university’s website and some misleading numbers were verified and corrected. In this case only numbers from after the PBES’s implementation were available, which meant that the previously determined time frame had to be deserted. The analysis, therefore, made use of numbers from 1999 to 2008. Furthermore, the format of those publications awarded points by the system was collected. Similarly, these numbers were obtained from the UoI website, but again for a different period. This time only three years had been processed in such a manner as to differentiate between ISI articles and other articles, that is, 2006-2008. This distinction was seen as an important element in the discussion relating to the equality of the system, and as a result this particular aspect of the bibliometric study only covered three years.

1.3 Summary

This chapter has discussed the choice of a theoretical framework, as well as the research strategies implemented in this study. Being an evaluative research, the appropriate theoretical construction within that field was determined. Having reviewed possible
theoretical approaches, the one offered by theory-driven evaluation was chosen on the basis of its flexibility, as well as its firm framework.

The framework was built around a programme impact theory of the PBES, where the aims of the study, as well as relevant variables, were described and discussed. The aims were threefold, that is, to examine the system’s publications and citations, the perceived impact of the system on research behaviour, and finally the system itself for explanatory purposes.

The research strategies, for answering the questions posed by the programme impact model, were then determined and discussed. For a sounder evaluation a triangulation of methods was decided upon, that is; interviews, survey implementation and bibliometric analysis. Those research strategies were discussed in relation to the possible limitations and alternatives, as well as further information about the research execution.
Chapter 2 - Literature Review

Given the aims of the study, this chapter reviews the current literature in the topic area of bibliometrics usage in research evaluation, and identifies the most relevant sources thereof. Firstly, bibliometric indicator are reviewed, as well as the possible appliance and limitations in their usage. Secondly, a discussion of the use of bibliometrics in research evaluation takes place, followed by a detailed discussion on three different systems of research funding and economic research encouragement via bibliometric means. Additionally, research policy at the UoI will be presented, along with the Icelandic system of performance based evaluation; the object of analysis in this dissertation. Lastly, a summary of the main points of the abovementioned topics closes the chapter.

2.1 Bibliometric Indicators – Appliance and Possible Issues

Before delving further into the realm of funding allocation via bibliometric evaluation systems, it is important to take a closer look at the indicators frequently used in evaluation of scholarly work. Bibliometric methods are, in general, made to measure “…counts of the frequencies with which events of specified types are observed to occur” (Borgman & Furner, 2002, p. 7). This idea is the core of the two most extensively used bibliometric methods in research evaluation; publication and citation analysis. This section will give a brief overview of these two indicators, their possible usage, as well as appliance. Furthermore, theoretical, technical and methodological issues applying to evaluations, which rest on these two indicators, will be examined.

2.1.1 Publication and Citation Analysis

The idea of monitoring scientific activity by bibliometric means is hardly a new invention. The earliest account of such dates back to 1917, when Cole & Eales examined comparative anatomy publications from 1543 to 1860. They did this by simply counting the number of journal articles and books, and then dividing them by country (Garg, 2003). Other
studies followed during the next years, but it was not until the emergence of the Science Citation Index in 1963, and subsequently the development of computerised databases, that this field of researched started to blossom (Garg, 2003).

This first exercise by Cole and Eales can be seen as a typical example of publication analysis, whereby scholarly publications are counted to construct an image of academic activity. Similarly, citation analysis, which focuses on linkage between scholarly works, counts citations for the purpose of monitoring scholarly communication and possible impact. Both of these methods can be used on different levels of aggregation, that is, to measure activity and impact of individual scholars, groups, institutions and departments, and even up to national or continental level. Additionally, the indicators can be used to evaluate specific fields (Borgman & Furner, 2002).

In order to be able to conduct a publication or citation analysis, one needs to have access to countable scholarly information of production and citations. This kind of information can be accessed via citation databases, such as; Thompson Reuters‘ WoS, Scopus and Google Scholar; more specified domain databases; or via institutions themselves, such as annual reports or institutional databases (Larsen & Ingwersen, 2009). The coverage of these different information portals will be discussed later in the technological section below.

2.1.2 The Theoretical Aspect

Even though bibliometrics can simply be viewed as a way of measurement used for various evaluations, this field of research also offers a theoretical debate in relation to citations. Although this particular dissertation is not able to extend its scope to encompass a thorough discussion of citation theory, the point must be made that various schools of thought have emerged on what citations actually measure. A concise introduction of both schools and
scholars, in relation to citation theory, is available in Moed’s work on citation analysis in research evaluation (2005).

Those disciplines, which have discussed and debated citations and citation behaviour, are those of physical sciences, which primarily focus on scholarly activity in quantitative terms, as well as the construction of indicators. Secondly, the social sciences have also studied citations, and consider them to be parts of social acts of researchers. As a result, citations are, in the role of social acts, able to cast light on how the science community evaluates research contributions in relation to significance (Moed, 2005).

Furthermore, the discipline of psychology has also made contributions to this debate, especially wondering what motivates academics to cite (Moed, 2005). In addition, historical approaches have emerged, which use citations to discern the development of scientific ideas and those who stand behind them (Moed, 2005). Last, but not least, are the approaches which belong to the disciplines of information and communication science. This field of science researches, for example, the communication of scholars and dissemination of information, (Moed, 2005), which can also be applied to citations and the analysis thereof.

These different approaches cast light on the diversity within the citation debate, as well as how the use of citation analysis within research evaluation can be problematic. At least the assumptions about what citations are measuring must be clearly defined. Hence, Moed (2005) stresses that further interpretation by the evaluator must be undertaken in order to clearly define what is being measured by counting citations.

2.1.3 The Technological Aspect

Conducting a comprehensive publication or a citation analysis requires the use of specific publication and citation platforms. Thompson Reuters’ WoS is perhaps most frequently used for such activities, as well as Scopus and Google Scholar, after their recent
emergence. The choice of platforms can be an important factor in relation to what kind of an analysis is to take place, as, for example, WoS is dominated by “hard sciences”, whereas Google Scholar incorporates also books and book chapters, which would suit arts and humanities or social sciences analyses better (Moed, 2005; Meho & Yang, 2007).

Despite the merit of being able to access an abundance of publication and citation information in one place, these platforms and their usage can introduce a series of potential technical problems. Van Raan (2005), pointed towards some of the technical issues regarding WoS. Firstly, van Raan reports the issue of lost citations. What happens is that an error occurs when matching the reference of published articles (citing articles) with formerly published articles (cited articles) – hence citations are lost. This could inevitably skew numbers, and as reported by Moed (2002), is certainly the case with 7% of non-matched references. Throw in other technical errors, such as wrongly spelled author names, mistyped volume or page numbers, to name a few, and you reach an error number of 25%-30% (Moed, 2002).

Another technical problem reported by van Raan, centres on the “cleaning up” of the raw data produced by the databases. Because of multiple spellings of university names and numerous differently named colleges and departments affiliated with the same university (especially the case with university hospitals), or even spelling conventions in relation to author names from various countries, it can become exceedingly difficult to access the “true” number of publication output by a singular institution or scholar (van Raan, 2005; Charon & Wauters, 2008). This specific issue would of course also be applicable to Scopus and Google Scholar.

Further technical problems have also been discussed by various authors. Derrick, Sturk, Haynes, Chapman, & Hall (2010) point to the importance of scholars to identify which database within the WoS is used as they differ, and could presumably make the repetition of
research difficult. Additionally, Weingart (2005) adds that because of the easy accessibility and user friendly interface of WoS, and now presumably also Scopus and Google Scholar, bibliometric methods have become a tool usable for “amateur” bibliometricians, who might not be aware of the assumptions that lie behind the numbers they are dealing with. Weingart (2005) also mentions ISI’s monopoly as a data provider, although that ceased to be the case with the launch of Scopus and Google Scholar in 2004 (Fingerman, 2004; Google, 2010).

2.1.4 The Methodological Aspect and Possible Pitfalls in Appliance

Having discussed the technical issues with using bibliometric indicators, it is pertinent also to consider methodological problems, which could arise should they not be highlighted or introduced to those undertaking evaluation procedures. Again, these issues are widely addressed by van Raan (2005), Ioannidis et al. (2007) and Moed (2005). The two former ones do so in relation institutional evaluation (Shanghai and Times rankings), whilst Moed gives a more general view on possible methodological issues when using bibliometric methods.

The first, and perhaps the most obvious issue, is the one of subject coverage and the differences in both publication and citation behaviour between disciplines. The “hard sciences” have a dominant coverage in journal based citation indices, whilst art and humanities and social sciences tend to communicate more through other means of publications, such as books (van Raan 2005; Moed 2005). In addition, the difference between disciplines can be identified in terms of the quantity of publication output, as some disciplines (such as medicine) produce more than others (Sandström & Sandström, 2009).

This is not only the case with different disciplines, but can also affect subfields within disciplines. When determining citation impact, Lazaridis (2010) points to the case of chemistry, a discipline with various subfields, which have different citation patterns. For example, organic chemistry collects a much larger proportion of citations than environmental
chemistry. As a result, the mere choice of focus of an institution’s research field becomes a
determining factor, making it difficult, if not at all impossible, to evaluate a whole institution
using simply journal articles as a form of measurement.

Furthermore, the field specific behaviour can also be viewed in the difference between
citation patterns of disciplines. These different characteristics become evident when
examining the time which takes the various fields to respond to new research in terms of
references (Moed et al., 1985). As a result, evaluators need to pay specific attention when
using citation in research evaluation, especially when working with a small citation window.

Another important methodological issue is the one of language. As pointed out by van
Leeuwen, Moed, Tijssen, Visser & van Raan (2001), Grupp, Schmoch & Hinze (2001), as
well as van Raan (2005), the fact that non-English journals included in the ISI citation index
tend to have a lower impact (presumably because of a narrower audience), the inclusion of
such journals in bibliometric research can skew the overall picture for some countries. As a
result, the impact measured for these articles is virtually none, although they do count as a
part of a publication output. In the study performed by van Leeuwen et al. (2001) this was
particularly seen to affect countries such as France, Germany and Switzerland. The case is
even more severe in relation to developing countries, which are heavily under-represented
within international bibliometric databases (Packer & Meneghini, 2007). Again, this could
prove a dangerous pitfall in research evaluation.

Yet another prominent issue in regard to methodology relates to the size and nature of
the institutions being evaluated. According to Ioannidis et al. (2007) it is necessary to account
for different sizes of institutions, since obviously larger institutions presumably produce more
work and receive more citations. These numbers will have to be normalised. Additionally, the
importance of a specific time frame when dealing with both publication and citation
measurements, and accountability for average numbers as well possible extremes, which might skew the curve within institutions, should be given a careful consideration by those conducting research evaluation (Ioannidis et al., 2007).

Last, but certainly not the least, is the problem of credit allocation. The task centres on the question of co-authorship, that is, how should one allocate credit to multiple authors; simply count the first one or divide credit amongst the whole group? In needs to be taken into account that various disciplines have different approaches in this matter, some listing the main contributor first or last, whilst others adhere to alphabetical order (Research Information Network, 2009). Similarly, it is worth pondering if top award winning researchers should be allocated credit as member of their present institution, or the institution he/she was affiliated when working on the groundbreaking research (Ioannidis et al., 2007). These are all matters that need to be worked out should one choose bibliometric methods to measure excellence.

2.2 Bibliometrics in Research Evaluation

The last decades have seen the importance and interest in research evaluation, as well as ranking of universities, grow extensively. The reason for this increased significance of research evaluation can be contributed to a growing demand for accountability, efficiency and effectiveness by funding bodies, such as governments (OECD, 1997; Abramo, D’Angelo and Caprasecca, 2008). The ultimate goal of “getting one’s money’s worth” has, therefore, been introduced into the research community. Hence, one can foresee that the current dip in the economy, and the subsequent diminishing of funds and capital used for research, will undoubtedly contribute to an additional strengthening of research evaluation on the whole.

As a result of the rising interest in economically centred research evaluation, parties at various levels in the community, such as institutions, departments and even national funding bodies have developed various systems of funding allocation based on a research evaluation.
Geuna & Martin (2003) offer an excellent overview and an international comparison of such evaluation systems in Europe and the Asia-Pacific region. Therein, as in other literature, it is noted that the main two methods used in research evaluation are bibliometrics and the more extensively used peer review (Geuna & Martin, 2003; Abramo, D’Angelo and Caprasecca, 2008), although the usage of bibliometrics seems to be on the rise (Sandström & Sandström, 2009; Schneider; 2009).

The reason for the increased interest in bibliometric within the field of evaluation research is not concrete; however one can imagine that much of it has to do with the fact that it is quick to perform, convenient and also cost effective (Lazaridis, 2010). Additionally, Weingart (2005) believes it could be contributed to disinterestedness and scepticism of peer review, as well as the fact that bibliometrics offer objective numbers and an abundance of them. Consequently, possible ulterior motives of citation would be neutralised.

2.3 Bibliometrically-Based Research Evaluation Systems

As the purpose of this review, and the focus of the overall research, centres around bibliometric indicators, this section of the review will cast its spotlight on research evaluation systems, which are based upon bibliometric indicators and in some form or another are intended to encourage, as well as evaluate, research in relation to funding. The ones chosen for a more detailed review here are the national system for university research funding in Norway, hereafter referred to as the Norwegian Model, Australia’s Research Quantum (RQ), and, to introduce the discussion of citations within such systems, Sandström & Sandström’s proposed system of research evaluation for Sweden. All of these systems are bibliometrically-based. The Icelandic system of research evaluation at the UoI will be introduced in a later section.
One of the research evaluation systems, which radically changed the way research was conducted, was the Australian RQ, introduced in Australia during the 1990s (Butler, 2002). The RQ was a part of dual funding system for research operated by the Australian government (Butler, 2008), which partially allocated funds for research and training to universities based on publication counts (Butler, 2002).

A more recently implemented bibliometric model for research funding was established in Norway. It was commissioned by the Norwegian government in 2002, and then implemented for budget allocation in 2006, after having gone through extensive development by the Norwegian Association of Higher Education (Schneider, 2009). The purpose of the system is twofold, that is, both to establish an extensive system of documentation of all academic publications in Norway - in order to use the information for research funding allocation by the Norwegian government, and to encourage research activity (The Norwegian Association of Higher Education Institutions [NAHEI], n.d.).

The proposal for the Swedish system, based on ideas from Ulf Sandström and Erik Sandström, closely resembles the Norwegian Model, as it is based on publication counts. It does, however, offer a new dimension with the introduction of citation usage, as well as a new approach to normalising publication counts by fields (Sandström & Sandström, 2009). These three aforementioned systems will now be analysed in relation to publications, citations, as well as possible effects on collaboration and equality.

2.3.1 Using Publication Counts and Determining Impact

Although information regarding the old Australian RQ is hard to come by, Linda Butler wrote numerous articles on the system’s function in encouraging research activity. Butler called attention to the soaring numbers of ISI publications from Australia in her article “Explaining Australia’s Increased Share of ISI Publications – the Effects of a Funding
Formula Based on Publication Counts” in 2003. The article strived to answer the question of the increased publications of up to 25%, at the same time Australia’s citation impact was falling.

Butler came to the conclusion that the reason was to be found within the university sector, where, as a result of the publication aspect of the RQ, researchers had changed their publication habits, in order to gain from the system. Hence, simply the publication counts could be translated into dollar amounts researchers could earn their institution.¹ As a result, the system created a research environment, where the publication of large amounts of articles was economically the most feasible behaviour, but not necessarily publications in the most prestigious publication channels.

At a first glance, the Norwegian policy makers seem to be heading down the same road as their Australian colleagues; placing the focus simply on research output. The Norwegian Model does, however, not only count publications and allocate funding accordingly, but also distinguishes between publications channels, offering extra research points for publications in outstanding journals, books etc. (NAHEI, n.d.). In this way the model tries to guide publications towards prestigious publication channels, thereby encouraging high level research without sacrificing quality for quantity.

¹ Butler provides the example of an article in refereed journal being worth just over $800 AUD, and a book
In order to encourage high level research, the Norwegian Model is made up of two levels. Level 1 is seen as incorporating all “normal” academic publications, whereas only leading and outstanding publication channels are allowed on level 2 (NAHEI, n.d.). The publication channels on level 2 are also subject to the 20% rule, that is; “At any one time publication channels on level 2 can account for a maximum of 20 per cent of the world’s publications within a specific field”, making level 2 flexible in relation to changes in publication behaviour (Schneider, 2009).

The Swedish proposal goes down a different road to determine and encourage the quality or impact of research output. The idea centres around counting citations, which have been field normalised beforehand (Sandström & Sandström, 2009). According to Sandström & Sandström (2009), this provides an even better picture of a research group’s level of research than the two tiered Norwegian way, as actual citations are measured instead of the impact of the journals, in which researchers try to publish their results.

As evident from the theoretical citation debate above, using citations in research evaluation can be a complex issue, and perhaps too underdeveloped to use in budget allocations. Furthermore, Schneider makes a strong case against such usage, acknowledging the fact that academic fields have inherent differences in terms of citation accumulation, making the evaluation of research less contemporary and not viable for a dynamic funding model (Schneider, 2009). How large a citation window the Swedish Model will end up using is unknown, but this added feature of impact or quality measurement could easily pose a problem.

<table>
<thead>
<tr>
<th>Publication Type</th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article in an ISSN-title</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Article in an ISBN-title</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>An ISBN-title</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2.3.1.1 – The two dimensional Norwegian Model.
2.3.2 Allocating Credit amongst Contributors – Possible Effects on Collaboration

When viewing the respective models in terms of their effect on collaboration, the designers of the Norwegian Model have aired their concerns about the possible side effects of using fractionalised counting of research points in cases of multiple authorships. The Committee sees this approach as a potential negative incentive for research collaboration (NAHEI, n.d.). The question of collaboration effect has, however, not been taken into thorough investigation within the Swedish proposal - at least such a debate has not been widely published.

It seems that no thorough investigation of the possible discouraging effects of the Norwegian Model on collaboration has taken place. Schneider (2009), however, argues that no empirical indications for such changes in publication behaviour are evident up to date. In terms of selecting another method of calculation it is of course easy to argue that applying whole counting, and thereby crediting each institution with full marks, would undoubtedly result in another RQ dilemma, as well as possible free riding collaborators. There is, therefore, no apparent cure for this invisible disease, that is, if it does exist. Nevertheless, some kind of collaboration incentive might be tied to the funding via other means. Despite the current positive outlook, it is, however, necessary to keep this point in mind, and to do an extensive investigation of this aspect of the model, especially now as time and experience have granted researchers with enough material for such an analysis.

2.3.2 Applying the Models to All Disciplines – Questions of Equality

When contemplating a system’s or a model’s capability to evaluate all research output equally, a number of questions become apparent. Firstly, there is the question of how the system deals with differences in terms of output quantity, as discussed above this may vary considerably between disciplines. Secondly, there is the dimension of coverage by the system, that is, how sufficiently does the system represent all relevant output from a specific
institution? Thirdly, there is the question of how research output, written on a specific local matter, and perhaps also in the researcher’s native language, is treated by the system, as these publications might not as easily gain an international status.

In terms of the first question, both the Norwegian Model and the Swedish System have considered the difference in the quantity of publication output by disciplines. The Norwegian Model has the normalisation inherent in its two level approach, as the 20% rule secures a constant number of publications within each field (Elleby & Ingwersen, 2010). The Swedish proposal does, however, approach this matter differently, opting for a field normalisation of all publication output via Waring distribution calculations (Sandström & Sandström, 2009). Both models seem, therefore, to be aware of this possible threat to equality within the system, although it must be noted that the Swedish method, as novel and interesting as it seems, might be at the price of the transparency of the system.

Secondly, when examining coverage of output, the Norwegian Model evaluates all academic publications, irrespective of disciplines, that is, if they present new insight, are presented so that the findings can be verified or used in a new research activity, are written in a language and distributed so that it becomes accessible to most researchers, and appear in an externally peer-reviewed publication channel (NAHEI, n.d.). As a result, the system grants points within three categories of publications; articles within ISSN and ISBN titles, as well as ISBN-titles (Schneider, 2009).

Judging from Sandström & Sandström’s (2009) article on the Swedish proposal, all publications are to be evaluated using an international publication and citation platform, such as the WoS. This of course makes one wonder about how this will affect the arts and humanities, as well as the social sciences, because of the known lack of data regarding these specific fields within the WoS (Moed, 2005). Instead of looking elsewhere for information, it
seems as if the Swedish Model proposes that those publications and citations within the two aforementioned fields, which are found in the WoS, be used as a yardstick for further calculations of the possible extent of the output and impact of the arts and social sciences (Sandström & Sandström, 2009).

Even though Sandström and Sandström believe the “proof to be in the pudding” (p. 246), estimating quantity of publications, and the citations gained there from, seems not only dubious, but could, apart from either over or underestimating the output from these fields, cause friction and bad moral within the Swedish research community. Should this be the road eventually taken by the Swedish government, Sandström & Sandström’s claim of the superiority of the Swedish Model on grounds of accurate presentation of the level of research via citations, can hardly be deemed valid.

Finally, how the systems have handled the matter of publications presented on specific nation related issues, perhaps also in the respective languages, is also a matter of equality. The creators of the Norwegian Model make a note of this issue, deeming publications published in Norway and in Norwegian to be bound for level 1, due to the lack of the international perspective (NAHEI, n.d.). The model’s purpose is, therefore, to encourage those using national publication channels to move to a wider audience with more internationally oriented publications (NAHEI, n.d.).

Despite being a valid point, this aspect of the system could also change research behaviour for researchers focusing on specific Norwegian topics. One could imagine that a researcher of the Norwegian language or culture might be more tempted to publish more in level 1 journals, as level 2 journals might be a far stretch for such an isolated topic. The other change might, however, be so that the same researcher mentioned above might move away
from her/his research focus, and start researching topics more acceptable in international level journals. As a result some subfields might suffer.

The Swedish proposal does not discuss this topic, but judging from the proposal those publications in Swedish or on a specific Swedish topic, would be evaluated as other publications, and presumably fall into the “pudding” category of predicted publications and citations. How this would affect research behaviour is hard to say, and would probably rest on the outcome of the estimation processes.

2.4 Research at the University of Iceland and the Icelandic Performance Based Evaluation System

Today the UoI’s research policy is determined by both by the Icelandic government and the university itself. Firstly, the UoI operates under the University Act no. 85 from 2008. The Act contains the following clause from the third article on the role of public universities:

A university tends to teaching, research, the seeking of knowledge and creation in the fields of science, academia, technology development or art. An education, provided by a university takes into account the needs of society and can be of academic or vocational nature. A university is the centre of knowledge and a part of an international education and science society.

(The Act on Public Universities, 2008)

The research role of the UoI is, therefore, grounded in the laws of the country, making it one of the foundations of the modern university.

Secondly, the Science and Technology Policy Council, which adheres to the Prime Minister’s Office, determines national aims in science and innovation. In the council’s most recent manifesto, which applies from 2010 to 2012, the focus is on the “three guiding lights”, that is, collaboration and sharing, quality and rewards, international research and innovation
(The Science and Technology Policy Council [STPC], 2010). The manifesto, therefore, encourages universities, as well as other research institutions, to increase cooperation and sharing in order to maximise efficiency and research outcomes; stresses that public funding for research be more connected to quality and outcome evaluation, and outlines the importance of the international angle in research and innovation (STPC, 2010).

Thirdly and lastly, the university itself also sets forth research objectives. These aims are more thoroughly pinned down and detailed than the above mentioned laws and policies. The long term goal, stated by the university’s current rector, Kristín Ingólfsdóttir, is for the university to become one of the top 100 universities in the world (UoI, 2006a). In order to achieve this goal the university aims at providing its students with outstanding teaching and support services, as well encourage outstanding research (UoI, 2006a). It is precisely in the last mentioned criteria that the reason for the creation of the current research evaluation system becomes apparent, as it can be seen as an encouraging component on the road towards outstanding research.
<table>
<thead>
<tr>
<th>Goals for Outstanding Research</th>
<th>Specific Steps in Goal no. II</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Quintuple the number of its yearly PhD graduations (graduate at least 65 doctoral candidates in the year 2011).</td>
<td>Number of papers published in international peer-reviewed journals is to increase by 100% by year-end 2011. The reward system for research will be revised to give greater weight to such publications. Special recognition will be given for papers published in world’s leading journals in each field of scholarship, such as <em>Nature</em> and <em>Science</em>. Special recognition will also be given for books published by highly respected international academic publishers. Changes are to take effect in 2007.</td>
</tr>
<tr>
<td>II. Increase research activity and quality of research – increase number of papers published in international peer-reviewed journals.</td>
<td>In evaluation of research publications, due recognition will be taken of the fact the University of Iceland plays a leading international role in research on Icelandic culture and society. Therefore, researchers in these fields are encouraged to publish in respected peer-reviewed Icelandic journals and books.</td>
</tr>
<tr>
<td>III. Increase organised collaboration with universities and university faculties overseas which are in the first rank in the world.</td>
<td>Professorial positions will be defined for highly-cited researchers. At least one such research will be appointed by 2011.</td>
</tr>
<tr>
<td>IV. Seek more funding from competitive funds.</td>
<td>Number of postdoctoral students is to double by 2011. Their facilities will be improved. The University of Iceland’s contribution to postdoctoral grants will be increased in stages over the whole period.</td>
</tr>
<tr>
<td>V. Greatly improve facilities for research and teaching, and increase access to on-line journals and databanks.</td>
<td>Eligibility for sabbaticals will be made contingent upon research activity. Rules to be issued by year-end 2006.</td>
</tr>
<tr>
<td>VI. Increase cross-disciplinary research.</td>
<td></td>
</tr>
<tr>
<td>VII. Promote innovation and links with research bodies, the economy and the regions.</td>
<td></td>
</tr>
</tbody>
</table>

(UoI, 2006a)
2.4.1 *The Icelandic Performance Based Evaluation System*

Writing a concise introduction of the Icelandic PBES proved more of a challenge than was expected. The reason for this was simply that almost no written material, except for the rules of the system and single report (Sigfúsdóttir et al., 2005), existed. As a result, it was necessary to consult people within the UoI environment or those in other ways connected to the system’s implementation, for further information. Various parties were contacted, most of whom agreed to answer a few questions regarding the PBES. The group of interviewees contained the initial creators of the system, current teachers and researchers at the UoI, staff of the Division of Science and Research at the UoI, as well as a government employee involved in the initial implementation. The information provided by those parties proved extremely useful in the production of this dissertation, which will attempt to provide a thorough discussion of the system, and hopefully fill in some missing knowledge blanks.

The PBES of the UoI, and other universities operated by the state, was originally launched in 1998. Its initial establishment contained the goal of increasing the salaries of university professors, by portraying an obvious and definite output of research, as well as simultaneously encouraging further research at the university. The system was, however, not modelled on any other pre-existing system, although ideas from surrounding countries were used as building blocks. The current version of the system, which has gone through some modifications in the last years, is from 2009.

As the Norwegian system, the PBES is grounded on publication counts, but differs from the way it affects the university staff. The PBES notably rewards individual researchers instead of whole institutions, and does so both in terms of salaries and career advancements. Firstly, the system operates as a distributor of bonuses from a communal bonus pot, basing its annual allocation on individually earned research points for publications. The amount available for distribution from the pot each year is determined by 12.5% of the salaries of staff
members. As a result, the pot has a fixed amount each year, so the value of research points in Icelandic kroner can differ, depending on how many points are earned each year - the more points the less they are worth.

In addition to affecting the bonus payments, the PBES also has an effect on basic salaries of staff members. When a person is employed by the university all material already published by that individual is counted as means of determining the position within the academic hierarchy (such as adjunct, lector, docent and professor). Furthermore, the system is used in relation to promotions, as a certain amount of points moves an individual up the aforementioned hierarchical ladder.

As mentioned above, the system is based on publication counts, as well as counts of various other activities believed to contribute to the furtherance of science, such as innovation activities and grants from competitive funds. The following categories are supposed to grant research points:
<table>
<thead>
<tr>
<th>Research Output</th>
<th>Points Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theses on post graduate level</td>
<td>(15-30 points)</td>
</tr>
<tr>
<td>Books</td>
<td>(10-100 points)</td>
</tr>
<tr>
<td>Book chapters</td>
<td>(0-20 points)</td>
</tr>
<tr>
<td>Academic articles</td>
<td>(5-20 points)</td>
</tr>
<tr>
<td>Articles in conference proceedings</td>
<td>(3-10 points)</td>
</tr>
<tr>
<td>Lectures and posters</td>
<td>(1-5 points)</td>
</tr>
<tr>
<td>Editorial work on academic publications</td>
<td>(3-20 points)</td>
</tr>
<tr>
<td>Reports, reviews and translations</td>
<td>(0-25 points)</td>
</tr>
<tr>
<td>Citations</td>
<td>(0.05 – 1 point per citation)</td>
</tr>
<tr>
<td>Curriculum design for preschools, primary and secondary schools</td>
<td>(0-10 points)</td>
</tr>
<tr>
<td>Innovation and knowledge transfer</td>
<td>(0-40 points)</td>
</tr>
<tr>
<td>Grants from competitive funds</td>
<td>(0-20 points/year)</td>
</tr>
</tbody>
</table>

Table 2.4.1.1 – Research output granted research points at the University of Iceland
The researchers at the UoI are, therefore, encouraged to publish, but the amount of research points available within each category (such as articles or books) is related to the perceived quality of the publication channel, in which the publication takes place. For example, when looking specifically at books, authors can receive 10 and up to 100 points for their publications, all depending on how prestigious the channel is perceived to be. This is also applicable to the majority of research activity under evaluation by the system, such as academic articles, book chapters, lectures and posters, as well as editorial work. A further breakdown of points rewarded for journal articles and book are visible in table 2.3 below. The full list of evaluands is available in appendix III.

| Peer-reviewed publications by the world’s most respected academic publishers (up to 100 points) | Article published in ISI-journal with high impact factor or in an A category journal in the European Reference Index for Humanities (ERIH) (20 points) |
| International peer-reviewed publications and national peer-reviewed publications with an international dimension (up to 75 points) | Other articles in ISI journals, B category articles (ERIH) or articles in journals that receive a grade I rating in journal survey (15 points) |
| Peer-reviewed publications (Icelandic or foreign) primarily aimed at the domestic academic community (up to 50 points) | C category articles (ERIH) and articles in journals that receive a grade II rating in journal survey (10 points) |
| Other books (up to 25 points) | Articles published in journals that receive a grade III rating in journal survey (5 points) |
| Republications (up to 10 points) | |

Table 2.4.1.2 – An example of how articles and books are granted different numbers of research points based on the publication channel and its diffusion capacities (The Evaluation System Committee [ESC], 2009).
Another aspect of the system worth a mention in this introduction is the fractionalised counting of research points in multiple authorship publications. The basic outline of these calculations is based on the following points:

- **Two authors**: $1.5 \times \text{points} / 2$
- **Three authors**: $1.8 \times \text{points} / 3$
- **Four or more authors**: $2.0 \times \text{points} / \text{number of authors}$

(ESC, 2009)

As a result, an article granting 15 points would provide each author with 11.25 points should the authors be two, nine points if they authors were three, and then two times the article’s worth divided by the number of authors should they be more than three.

In addition to these calculations, additional points are awarded on the basis of the frequency of multi-author publications. If a researcher publishes four multi-author publications or less during the year of evaluation, she/he is awarded additional points for one work (the work awarded the most number of points after calculation). For example, if an article worth 15 points is written by three authors, each author would receive three points. The difference between the total points awarded and the ones received by each author is, therefore, 12. This number is then divided by two, resulting in six additional points. Furthermore, should an author publish more than four pieces of multi-author work annually, the same principle of calculation is used for adding additional points for two works instead of one (ESC, 2009).

Lastly, it must be noted that a specific evaluation committee, containing representatives from all academic schools, is granted the authority to depart from the framework of evaluation and award more points to those publications perceived to be of such deserving. In addition, authors can have their work re-evaluated by the same committee, every
five years, if they believe the work to have had an extreme impact or attention in the scientific community (ESC, 2009).

2.5 Summary

Quite an extensively used indicator of research output and impact is the one pertaining to bibliometrics. Its increasing popularity can largely be explained by its cost effectiveness and ease of use, although many have commented on the various shortcomings and problems related to this method of measurement. These range from theoretical citation interpretations, technical issues of the various publication and citation platforms, to issues of a methodological nature, such as differences in publication and citation patterns within fields/subfields and language biases, as well as issues of normalisation and credit allocation.

A greater interest in research evaluation has been noted in the last decade as the demand for accountability, efficiency and effectiveness by funding bodies has grown. A number of universities now use, or have used, bibliometrically-based systems for research evaluation, such as Australia, Norway, Sweden and Iceland, although they do so differently. All system are or have been used to count publications, but the difference can be perceived when viewing how the publications are valued, how the system tries to manipulate publication behaviour towards prestigious publication channels, how it affects collaboration and equality, and last, but not least, how the funds earned by the research point collection are allocated.

The chapter ended with a short introduction of the Icelandic PBES, where the system was described as a means of explaining its intended function. The next chapters will present the results of the two studies carried out, that is, the bibliometric analysis as well as the questionnaire results. This information will hopefully grant a more thorough picture of how the system is working in terms of bibliometric numbers and the experience or views of main stakeholder – the academic staff.
Chapter 3 – Bibliometric Analysis Results

This section will outline the results of the bibliometric analysis, as it was laid out in the methodology chapter. This analysis includes a comparative view of the six institutions under examination; UoI, University of Bergen (UoB), University of Tromsø (UoT), University of Lund (UoL), University of Uppsala (UoU) and University of Copenhagen (UoC). Firstly, a two layered publication analysis including the total publication numbers and total publication numbers by subject categories is presented, followed by numbers reflecting the universities’ collaboration. Subsequently, a comparative citation analysis is introduced. The section then ends with an examination of research points rewarded to academic schools and departments of the UoI, and the publication format contributing to the points.

3.1 Publication Analysis

The first step of the bibliometric analysis was to determine the output of the six universities within the time frame of 1994 to 2002. As could be expected the total publications ranged considerably, as the size of the universities under comparison vary greatly. The chart below outlines the overall raw numbers for publication output and the development during the period. This raw approach only offers a very course analysis of the overall trends, but a few general assumptions can, however, be made. For example, it is evident that publications seem to be slightly on the rise for all universities, although the output from the two Swedish universities seems to be more fluctuating in nature during the period.
For the purpose of a more thorough comparison, it, however, became apparent that some sort of a normalisation or further processing of the numbers was imperative. As discussed in the methodology chapter, ways of normalisations were contemplated, and the most appropriate way in this instance, index calculations were applied. This allowed for a clearer view of the development of publication output from the six institutions. The results of these calculations can be viewed in chart 3.1.2 below.
When viewing the indexed numbers of growth from all the universities, the following points of interest become apparent:

- **UoT has the highest rise in publications, and although it experiences two instances of decrease during the period, its growth in publication output still remains significantly larger than of the other universities.**

- **The trend of publication output for all universities can be described as leaning towards growth. All universities do, however, experience a period of either one or two years where publications decrease from the year before.**

- **Four out of six universities experience a decrease in publication output in 1999.**

- **All universities have a growing publication output in 1995, 2001 and 2002.**

- **The UoI is the only university which encounters a decrease below the initial 100 mark established in 1994.**
• The UoI experiences a steep increase in publication output in 1998, followed by a similar fall the next year. Thereafter, publications can be seen as having a steady growth.

In order to investigate the publication output of the six universities even further, the total publication numbers were split up into the Noria-Net 14 subject categories; agriculture, art, biology, biomedicine, chemistry, clinical medicine, engineering, geosciences, information and computer technology (ICT), material sciences, mathematics, physics, social sciences and multidisciplinary sciences. This way growth in specific fields could be examined. In addition, the top five publication categories were determined, along with the top five categories in relation to percentage of growth, for all universities. Average percentage of growth was also established.

<table>
<thead>
<tr>
<th>Iceland</th>
<th>Bergen</th>
<th>Tromsø</th>
<th>Lund</th>
<th>Uppsala</th>
<th>Copenhagen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Med 27%</td>
<td>Clinical Med 29%</td>
<td>Clinical Med 28%</td>
<td>Biomedicine 24%</td>
<td>Clinical Med 23%</td>
<td>Clinical Med 34%</td>
</tr>
<tr>
<td>Biomedicine 16%</td>
<td>Biomedicine 16%</td>
<td>Biomedicine 18%</td>
<td>Clinical Med 24%</td>
<td>Biomedicine 19%</td>
<td>Biomedicine 27%</td>
</tr>
<tr>
<td>Geosciences 11%</td>
<td>Geosciences 10%</td>
<td>Biology 13%</td>
<td>Chemistry 12%</td>
<td>Physics 15%</td>
<td>Biology 8%</td>
</tr>
<tr>
<td>Physics 9%</td>
<td>Biology 9%</td>
<td>Geosciences 10%</td>
<td>Physics 11%</td>
<td>Chemistry 10%</td>
<td>Physics 8%</td>
</tr>
<tr>
<td>Social Sciences 9%</td>
<td>Social Sciences 7%</td>
<td>Agriculture 8%</td>
<td>Biology 6%</td>
<td>Biology 7%</td>
<td>Geosciences 5%</td>
</tr>
</tbody>
</table>

Table 3.1.1 The top five subject categories for each of the six universities determined by proportion of total publication output 1994-2002.

The first step was to determine the top five subject categories in relation to publication output from 1994 to 2002 (see table 3.1.1). The results gave a rather uniform image of the research activities within all six universities. In all but one instance clinical medicine ranks number one with the most published articles, followed by biomedicine at number two. Those two disciplines switch places when it comes to UoL, but the difference is only nine
publications. Furthermore, five universities have biology on their top five, and four of them physics and geosciences. In addition, chemistry and social sciences can be identified in two universities out of six.

<table>
<thead>
<tr>
<th>Iceland</th>
<th>Bergen</th>
<th>Tromsø</th>
<th>Lund</th>
<th>Uppsala</th>
<th>Copenhagen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material Sciences 500%</td>
<td>Multidisc. Sciences 467%</td>
<td>ICT 350%</td>
<td>Material Sciences 233%</td>
<td>Mathematics 121%</td>
<td>ICT 281%</td>
</tr>
<tr>
<td>Biology 325%</td>
<td>Geosciences 104%</td>
<td>Art 220%</td>
<td>ICT 158%</td>
<td>Material Sciences 114%</td>
<td>Geosciences 131%</td>
</tr>
<tr>
<td>Agriculture 300%</td>
<td>Physics 93%</td>
<td>Social Sciences 215%</td>
<td>Mathematics 100%</td>
<td>Agriculture 77%</td>
<td>Mathematics 109%</td>
</tr>
<tr>
<td>Geosciences 238%</td>
<td>Agriculture 81%</td>
<td>Engineering 167%</td>
<td>Biology 79%</td>
<td>Biology 54%</td>
<td>Social Sciences 107%</td>
</tr>
<tr>
<td>ICT 133%</td>
<td>Biology 51%</td>
<td>Agriculture 146%</td>
<td>Physics 72%</td>
<td>Physics 43%</td>
<td>Biology 94%</td>
</tr>
</tbody>
</table>

Table 3.1.2 The top five ranking subject categories of each university in relation to growth.

Moving on to the examination of growth within specific subject categories, the subject areas are not as uniform as was evident in the comparison above (see table 3.1.2). Five out of six universities do, however, experiencing growth in biology, four out of six in agriculture and ICT, and three out of six in geosciences, physics and mathematics. Surprisingly, the highest ranking categories in publication output do not reach the top five in any of the universities, that is, clinical medicine and biomedicine. The reason for this might simply be that these specific categories have a rather steady growth during the period.

Lastly, publication numbers for each category were summed up to represent the four years before and after the implementation of the PBES. For the purpose of comparison the increase or decrease in growth between these two periods was established. Average was also calculated for the same comparative purpose.
When viewing those numbers it becomes apparent that in only 6 categories out of 14, the UoI exceeds beyond the average in growth. The greatest difference is within the category of social sciences, where the UoI is 18% above the growth average when comparing the two periods. The other five categories, that is, physics, engineering, clinical medicine, chemistry and biomedicine have a growth rate above the average ranging from 5% to 16%.

In other categories, the UoI is, however, below the calculated average, with the categories of multidisciplinary sciences, mathematics and material sciences experiencing a tremendous decrease in publication growth. In this two period comparison mathematics have a 37% decrease in publications, multidisciplinary sciences 66%, and material sciences with the largest decrease of 108%.
3.2 Collaboration

As has been described in the previous methodological chapter, the level of cooperation, of the six universities in question, was determined by counting those publications between 1994 and 2002, which were not purely national. As a result, the total publications per year, which were found to include foreign collaboration in each of the universities, were divided by the annual total publication counts for each institution during the investigative period. This part of the analysis revealed the following main trends in collaboration patterns of the universities:

- The general trend for all universities is a rather steady growth in collaboration from 1994 to 2002. The collaboration numbers for the UoI are, however, more fluctuating, which can largely be credited to the small amount of publications used to establish the trend line.
- All universities experience a decrease in collaboration at some point, apart from the UoC, which has a steady upward growth from 31% to 46% during the period.
Most of the universities tend to stay close to the average, apart from the UoI, which has a rather higher proportion of collaboration than the other institutions.

In addition, the UoC stays slightly below the average during the period.

The next chart demonstrates the proportional average collaboration for each university for the 1994-2002 period. UoI sits on top with an average value of 52% whilst UoC has the lowest value of 38%. The other four universities have a very similar values ranging from 41% to 44%, which come closest to the average value of 43%.

![Chart 3.2.2 The average proportion of collaboration – all universities.](chart)

When taking a closer look at the status of the UoI in relation to the average of all the universities, the two are not drastically far apart.
As stated before, the main difference can be seen in the more fluctuating nature of the UoI average due to its small sample size. In addition, the UoI numbers are slightly higher than the total average, coming closest to the average line in 1995 and 2002. Three distinct increases can be identified from 1995 to 1996, 1998 to 1999 and then 2000 to 2001. The highest peak of collaboration comes in 1999 with a proportional jump in collaboration by 13%.

Finally, the growth of collaboration was determined from two periods, that is from 1994 to 1997 and then from 1998 to 2001. Those two periods represent four years prior and four years after the implementation of the PBES. The growth of collaboration at the UoI, from one period to another, amounts to 47%, which is 8% above the average of 39%. The highest ranking university in this respect is Tromsø, with a growth of 60%, which is double the growth of the lowest ranking one, Lund.
3.3 Citation Analysis

The next step taken in the bibliometric analysis concerned the citation impact of the publication output discussed above. Firstly, the sum of all citations obtained by institutions’ publications, from their time of publication to December 2010, was established using the citation analysis function within the WoS.
Chart 3.3.1 Sum of times cited for publication output 1994-2002 – all universities.

As is evident from chart 3.3.1 above, comparison between such differently sized institutions with different research emphasis became troublesome. In order to create a more comparable image a field normalised score for each of the institutions was calculated, as described in the methodology chapter. This was done by comparing the citation impact of all subject categories in each institution to the worldwide impact of citations in the same categories. The results can be viewed in chart 3.3.2 below.
Chart 3.3.2 The chart shows the Weighted Field Crown Indicator for each institution. As an example, the value of 1.2 represents a citation impact 20% above the worldwide score.

When viewing the citation trend for each university in 1994 to 2002, the following parts of interest become apparent:

- **Two universities remain above the average throughout the period. Those are the two Swedish universities. Only UoT remains below the average during the nine year time frame.**
- **All universities have a citation impact which exceeds the world baseline. The only anomaly is the UoI in 1998 with an indicator of 0.95.**
- **The UoU has a sudden increase in citation impact in 2001, but drops down to the average a year later.**
Another part of the citation analysis cantered on the growth of citation impact throughout the nine year period (see chart 3.3.3). It is evident that half of the universities in question exceed the average growth of 38%. The UoC does, however, come close with a growth value of 32%, whereas UoU lags approximately 10% behind with a value of 17%. The least growth in publication impact takes place at the UoI, which only has a 5% growth from 1994 to 2002.

The final part of the citation analysis encompasses the differences in citation impact averages before and after the implementation of the PBES. The results are visible in chart 3.3.4 below:
The most interesting aspect of this particular analysis is the fact that all universities show a growth in citation impact in the before period of 1994-1997 to the after implementation period of 1998-2001, except the UoI. Furthermore, when viewing the average of averages from the two periods, it becomes apparent that the UoI in fact exceeds the average of 1.31 in the before period with a value of 1.46, but has the second lowest average for the latter period of 1.27, when the average for the whole groups reaches 1.51.

### 3.4 Research Points and Publication Format

The last step in the bibliometric analysis concerns the question of equality, and is focused on the research points granted by the PBES, and the format of the publications laid out for evaluation. The Nordic comparative aspect of the other analyses is, therefore, not present here. In addition, as the research points were only granted after the implementation of the PBES in 1998, the time period under investigation is brought forward by placing the starting point at 1999 and then ending in 2008.
Firstly, research points for each academic department were collected from the UoI, and then summed up to represent each of the four academic schools under examination. Out of the four schools the School of Social Sciences was rewarded the most number of points during the period, or 27.578 points. The School of Engineering and Natural Sciences comes next with 27.213 points, followed by the School of Health Sciences with 24.910 points, and then lastly, the School of Humanities with 24.728 points. The difference between the highest school and the lowest is 2.850 points, or 12%. 

Chart 3.4.1 Total research points awarded to academic schools at the UoI 1999-2008.
Next the distribution of points over the period of 1999 to 2008 was examined. When looking at the School of Engineering and Natural Sciences the distribution of points follows the average line quite extensively, with a few drops and rises above the average, reaching its highest value in 2006. The School of Health Sciences starts off in 1999 with the highest proportion of points awarded that year, and then grows slightly during the next two years. In 2002 there is, however, a drop in numbers, which reaches it bottom value in 2003. Thereafter, the numbers of research points gradually increase, ending with a steep rise from 2006 to 2007, where the end value reaches slightly above the average.

Subsequently, the distribution for the School of Humanities was investigated. The number of points starts out at the exact location of the average, then slightly falling in 2000, and then rising steeply to its highest value in 2001, around 19% above the average. The points, however, for 2002 suffer a decrease which brings the amount of research points down
to the average again, but then they gradually grow slowly just below the average and end in a slight decrease.

Lastly, the School of Social Sciences starts out with the lowest number of points in 1999, but then rises to the average in the subsequent year. Nevertheless, a decrease takes place again in the next two years, reaching the bottom value in 2002. Thereafter, the social sciences have a great increase in numbers, rising steeply and towering 21% above the average, and 65% above the lowest value of the School of Health Sciences in 2004. The numbers then gradually decrease, finally reaching the average in the last year of measurement.

The last part of the analysis centres on the format of the publications awarded research points. This approach was taken in order to try to cast some further light on the possible inequality, and explain the difference in research points. Numbers, which differentiate between ISI articles and other peer-reviewed articles, were only available for the three year period of 2006-2008, resulting in a further shortening of the investigative period.

Chart 3.4.3 Form of publications awarded research points in 2006.
When viewing the numbers for these three years, the publication patterns for each academic school become apparent. Unsurprisingly, the School of Health Sciences and Engineering and Nature Sciences have the biggest portion of its output in ISI publications, ranging from 31% to 46%. Health Sciences’ second largest category of publication is in lectures or talks (21%-29%), followed by other (19%-22%), and other peer-reviewed articles (7%-12%). Finally, the minority of the publications can be found in books chapters (3%-9%) and books (0%-1%).

Chart 3.4.4 Form of publications awarded research points in 2007.

The School of Engineering and Nature Sciences has a slightly different ranking of publication formats from Health Sciences when the subsequent categories are viewed. The category, which includes the second largest number of publications, changes from book chapters (23%) in 2006 to lectures/talks (21%) in 2007, only to switch back to book chapters (24%) in the next year. Other material has a constant rank as number three (14%-16%).
followed by other peer-reviewed articles (5%-10%) and then a mini-proportion in books (0%-2%).

The trends take a different turn when viewing the numbers for the School of Humanities and the School of Social Sciences. Both of these schools have most of their work published in book chapters, or 24% to 31%. The second largest publication channel for the social sciences is lectures/talks (21%-25%), but the third place switches between other material (14%-16%) and other peer-reviewed articles (13%-22%). Subsequently, publication output is present in books (9%-12%), and then finally in ISI-publications (5%-7%).

Chart 3.4.5 Form of publications awarded research points in 2008.

The patterns for the humanities differ slightly, and have a more fluctuating nature. The second largest publication category switches from other peer-reviewed articles (27%) in 2006 to other material and other peer-reviewed journals (22%) the next year, only to change again to the other publications (23%) in 2008. The third rank goes to lectures/talks (21%-22%), and the forth to other material (15%) in 2006, lectures/talks (21%) in 2007 and other peer-
reviewed articles (16%) during the last year. The publication channels representing the minority of publications have a steadier rank, with books (6%-11%) coming in second last, and then finally ISI with the least amount of publications (1%-5%).

3.6 Summary

The results of the bibliometric analysis conducted within this study have been outlined in this chapter. The comparative analysis focused on publications, collaboration and citations in the six universities under examination. Subsequently, a specific investigation of the UoI’s research points, and the format of publications awarded those points, took place at the end of the chapter.

When viewing the outcome from the publication analysis, the most striking result is the significantly larger growth in publications during the period at UoT. Furthermore, the collaboration analysis revealed a number of interesting points, such as the slightly higher proportion of collaboration at the UoI in comparison to the other universities. In addition, the citation analysis revealed a decreasing trend of the citation impact at the UoI while publication numbers are on the rise. This particular point also needs further examination.

Finally, matters relating to the PBES of the UoI were also examined. This part of the analysis revealed a possible effect of the PBES on publication output in certain subject categories, such as multidisciplinary sciences, mathematics and material sciences. Additionally, the possible effect of the system on citation impact will be discussed in chapter five.
Chapter 4 – Questionnaire Results

This section will outline the results of the questionnaire presented to academic personnel at the UoI. The answers were processed using the SPSS software, which provided various methods to work with the data, both descriptive and analytical. The analytical tests conducted on the data gathered from the survey were one-way ANOVAs, independent sample t-tests, as well as relevant post-hoc tests.

The choice of analytical tests was based on recommendations obtained from the author of Discovering statistics using SPSS: (and sex, drugs and rock'n'roll), Dr. Andy Field at the University of Sussex. The one-way ANOVAs and t-tests provided the opportunity to detect differences in mean scores of the demographic groups on the Likert scale questions (Field, 2008a). As no concrete hypotheses were made beforehand as to the possible differences in answers between the groups, post hoc tests were conducted instead of a planned comparison. As a result, the post hoc tests were able to provide a clear answer as to which groups the significant difference applied (Field, 2008b). The post hoc tests chosen were those of Tukey and REGWQ, as both have shown good degree of power and type one error rate control (Field, 2008b). Additionally, the Games-Howell test was performed as a precaution procedure, in case the population variance might not be equal (2008b).

As a part of the one-way ANOVA analysis, Levene’s test was also conduct in order to discover if the variance of the groups were indeed equal. Should the assumption be false, SPSS was able to provide two other variants of the F-ratio, that is, Brown-Forsythe F and the Welch F (Field, 2008c). In cases where the assumption of homogeneity was broken, those two

The calculated means have the following meaning:
1: Strongly agree/Very satisfied. 2: Somewhat agree/Somewhat satisfied. 3: Neither agree nor disagree/Neither satisfied nor dissatisfied. 4: Somewhat disagree/Somewhat dissatisfied. 5: Strongly disagree/Very dissatisfied.
F-ratios were reported, and the post hoc test of Games-Howell performed instead of Tukey and REGWQ.3

4.1 Population and Response Rates

The population was determined by collecting e-mail addresses of all academic staff at the university's website. This initial selection was composed of 782 possible respondents, which was then reduced to 776, after the elimination of inactive e-mail addresses. Out of those 776, 241 responded to the survey, 191 of which were complete answers, and 50 in the state of started but not completed. As a result the initial response rate of the survey was 31%. Lastly, 379 viewed the survey without taking part.

When processing the survey responses, additional information from one of the interviewees in chapter two revealed that not all survey respondents were actively being evaluated by the UoI’s PBES. This group included the School of Education, formerly a separate institution of education science; Iceland University of Education (IUE), which merged with the UoI in 2008 (Sameining Háskóla Íslands og Kennaraháskóla Íslands, 2008). When operating as an independent university, the IUE had its own evaluation system, according to which education science personnel have been evaluated since the merger.

As this was seen as introducing the possibility of the answers reflecting the views and experiences of two different systems, the answers from those identifying themselves with any one of the three departments belonging to the School of Education, were removed. As a result responses were reduced to 199, 161 of which were complete answers and 38 from respondents starting but not completing. Therefore, the new response rate amounted to 26%.

3 It must be noted that when conducting the one-way ANOVAs on age groups the youngest group of 21-30 was removed. This was done because the group only contained one individual and, as a result, presented possible skewness in results.
In addition, it must be noted, that all “started but not completed” responses were viewed individually to determine if they should be kept within the dataset. All responses from this category, which were either completely empty, or had only demographic answers were eliminated before the data processing took place.

Chart 4.1.1 Population and response rate.

4.2 Demographic Questions

In the first section of the questionnaire respondents were asked to answer a set of demographic questions concerning their status within the university, which department they belonged to, the percentage of their research obligation, and then finally gender and age. The largest group of respondents consisted of professors (101), followed by docents (48), then lectors (25) and adjuncts (9). Fifteen respondents could not be classified in to the four aforementioned categories, but identified themselves as being scientists, academics, researchers, clinical lectors or professors emeritus.
Respondents were then asked to identify with which department of the university they were affiliated. The top five departments were as follows: Medicine (22%), Physical Sciences (10%), Life and Environmental Sciences (8%), Social and Human Sciences (7%), and Icelandic and Comparative Cultural Studies (6%). A chart depicting response rate of all departments can be found in appendix II. Only one department did not have any respondents, this was the department of Electrical and Computer Engineering. Subsequently, the departments were grouped together, as to represent the four academic schools of the university (apart from the School of Education), which yielded the following distribution: Health Sciences (35%), Engineering and Life Sciences (28%), Social Sciences (20%) and Humanities (13%).

\[N \text{ stands for the numbers of respondents who gave answers to the respective question.}\]
The next demographic question centred on research obligation. The purpose of including this particular question was to see how great a portion of a respondent's work was focused on pure research activities. As a result it could provide certain clues when crossed with other subsequent questions, as to the possible importance of this particular variable. The responses showed that 10 respondents had none research obligation, the majority, or 141 individuals reported a 40% research obligation, 18 had a 60% obligation and only 3 respondents claimed 80%. There were, however, 25 individuals, who chose the option marked "other", and reported percentages ranging from 8% to 100%, the most common values being 31% (4), 50% (4) and 100% (4). Interestingly, there were additionally three respondents who did not know, or were not sure of, the percentage of their research obligation.
Subsequently, participants were asked to identify gender and then age. Male respondents were 135 (68%), and female 57 (29%). When the population was determined, the gender was also noted – resulting in a 57% male and 43% female ratio. It can, therefore, be stated that 31% of all males in the population, as well as 17% of females responded to the survey. In relation to age, the largest group of the participants were aged 51-60, or 77 (39%). Other age groups were as follows: 21-30 had only one respondent, 27 (14%) respondents were aged 31-40, the 41-50 group had 56 (28%) respondents, and finally 37 (19%) were aged 61 or older.
Chart 4.2.4 Respondents' gender. N=185.

Chart 4.2.5 Respondents' age. N=199.
4.3 Satisfaction with the Performance-Based Evaluation System

The first question relating to the respondents’ opinions was their overall view of the system, that is, if they were satisfied with the current PBES. This first question set the tone for the forthcoming questions, establishing the basic view of the academic personnel. The split views on the system became apparent in this particular question, where 37% were either very or somewhat satisfied with the system, and 40% were either very or somewhat dissatisfied. The most common category of “somewhat satisfied” received 35% of responses. In addition, it should be noted that a much larger percentage of respondents leaned towards the very dissatisfied option (20%), than those claiming to be very satisfied with the system (2%).

When comparing the different demographic groups in their view on the PBES a significant effect was found of academic schools on overall satisfaction, $F(3,172) = 4.428$, $p < .005$. The two post hoc tests, Tukey and Games-Howell, indicated a difference between social sciences and health sciences on the one hand ($p < .034$ and $p < .038$), and health sciences and humanities on the other ($p < .010$ and $p < .026$). Additionally, the Tukey and REGWQ tests...
revealed the means for these groups to be as follows; health sciences 3.62; humanities 2.71 and social sciences 2.95.

4.4 Publication, Quality and Citations

In a subsequent part of the survey respondents were asked on their views on publication activity, portrayal of quality of publications, and the use of citations in the evaluation process. When looking specifically at research activity and publications, there seems to be an overall consensus amongst staff members that the PBES has an increasing effect on publication behaviour. A large majority of respondents, or 74%, believed the system to be a pertinent factor in the increase of publications. Another majority of 56% also saw the system as increasing research activity. Finally, 40% of respondents experienced the system as providing a correct image of their department’s publication output, whereas 26% somewhat or strongly disagreed with that statement. Non-responses in these three questions ranged from 16-17%.

![Chart 4.4.1 Respondents' views on statements regarding publication and research activity.](image)
Shifting the focus onto the questions of quality and citations, the aforementioned consensus in regards to increase in publications, is no longer apparent. About 30% of respondents believed the system to increase quality of publications, whereas 36% disagreed with that statement. Similarly, a quarter saw the system providing a correct image in relation to the quality of their department’s publication output. That statement was, however, disagreed upon by 39% of the survey participants. Lastly, when asked if citations should be used in a purposeful manner within the evaluation process, respondents seemed to lean towards taking up an active usage of citations within the PBES. Around 38% strongly or somewhat agreed to that suggestion, whereas 26% either somewhat or strongly disagreed. This leaves out about a fifth which neither agreed nor disagreed. Non-responses in these three questions ranged from 17-19%.

Chart 4.4.2 Respondents’ view on statements regarding the quality of publications and use of citations.

Conducting a one-way ANOVA revealed a significant effect of the position of respondents on their view on the PBES’s ability to provide a correct picture of their
department’s publications, $F(4, 161) = 2,759, p < .030$. The assumption of homogeneity was, however, violated; therefore, the Brown-Forsythe and Welch $F$-ratio were reported. The significant effect of position on the abovementioned statement was still present with $F(4, 31,128) = 4,459, p < .006$ (Welch) and $F(4, 84,355) = 3,774, p < .007$ (Brown-Forsythe).

As a result, the appropriate post hoc test of Games-Howell was performed, which revealed a significant difference between the positions of a docent and adjunct and docent and those in the group “other” ($p < .046$ and $p < .021$). In addition, the means for these three groups were reported by Tukey and REGWQ; adjuncts had a mean of 2,00, the group of “others” 2,33 and the docents 3,18.

In addition, age was revealed as having a significant effect on the same statement, $F(3, 160) = 3,962, p < .009$. Both Tukey and Games-Howell post hoc tests indicated that the difference was between the age groups of 41-50 and 51-60; Tukey $p < .024$ and Games-Howell $p < .036$. Furthermore, the Tukey and REGWQ test showed the means of these two groups to be 2,55 (51-60 years) and 3,20 (41-50 years).

Observing the views on quality of publications, age was perceived as a significant factor again. Asked if the PBES had a part in increasing quality in publications, age had the following significance: $F(3, 156) = 2,747, p < .045$. Both post hoc test, Tukey and Games-Howell, revealed the difference to be between those aged 31-40 years and the group of 60+ ($p < .033$ and $p < .031$). Lastly, Tukey and REGWQ disclosed the means of the groups to be 2,62 for the 60 years and older group and 3,65 for those aged 31-40.

Additionally, the one-way ANOVA revealed a significant effect of the age of respondents on their view on the PBES’s ability to provide a correct image of the quality in regards to their department’s publications, $F(3, 160) = 3,465, p < .018$. Again, both Tukey and Games-Howell indicated that the difference was to be found between those aged 41-50 and
60+ (p < .035 and p < .034). Furthermore, Tukey and REGWQ showed the means for these two groups: 2.90 for 60+ and 3.67 for those aged 31-40.

The ANOVA test also revealed a significant effect of age on research activity, \( F(3, 161) = 4.117 \) (p < .008). The assumption of homogeneity was, however, violated; therefore, the Brown-Forsythe and Welch F-ratio were reported. The significant effect of position on the abovementioned statement was still present with \( F(3, 70,852) = 4.100 \), p < .010 (Welch) and \( F(3, 126,075) = 4.352 \), p < .006 (Brown-Forsythe). When performing the post hoc tests of Tukey and Games-Howell, the former one indicated a difference between the age groups of 41-50 and 51-60 on the one hand (p < .040), and 41-50 and 60+ on the other (p < .008). The latter test did, however, only report a difference between those aged 41-50 and 60+ (p < .006). In addition, Tukey and REGWQ reported the means for each group, the 60+ had a 1.87 mean, the 51-60 group a 2.13 and those aged 41-50 had a mean of 2.71.

Furthermore, the views on PBES’ effect on research activity were also found to differentiate between academic schools. As a result, significant effect of academic schools on research activity was revealed, \( F(3, 159) = 3.655 \), p < .014. The assumption of homogeneity was, however, violated; therefore, the Brown-Forsythe and Welch F-ratio were reported. The significant effect of age on the abovementioned statement was still present with \( F(3, 66,441) = 3.473 \), p < .021 (Welch) and \( F(3, 90,982) = 3.572 \), p < .017 (Brown-Forsythe).

Both Tukey and Games-Howell tests reported a significant difference between academic schools in relation to the aforementioned statement. The difference was found to be between engineering and nature sciences and health sciences (p < .010 on both tests). Lastly, means of those groups were also reported, showing a 1.98 mean for engineering and nature sciences and a 2.69 mean for health sciences.
In addition, a significant effect was reported of academic schools on the views regarding the PBEB’s purposeful usage of citations in evaluation, $F(3, 157) = 6.350, p < .000$. The Tukey and Games-Howell post hoc tests indicated differences between social sciences and engineering and nature sciences ($p < .019$ and $.030$), humanities and health sciences ($p < .015$ and $p < .016$) and social sciences and health sciences ($p < .003$ and $p < .008$). The Games-Howell did, furthermore, report a difference between engineering and nature sciences and humanities ($p < .049$). Finally, Tukey and REGWQ showed the means of the groups in question; health sciences 2.44; engineering and nature sciences 2.58; social sciences 3.42 and humanities 3.43.

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>Dissatisfied</th>
<th>Sig. Level</th>
<th>New System</th>
<th>Not a New System</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increases Publication</td>
<td>1.38</td>
<td>1.82</td>
<td>.003</td>
<td>1.41</td>
<td>1.76</td>
<td>.049</td>
</tr>
<tr>
<td>Provides a Correct Image of Dep. Pub. Output</td>
<td>2.26</td>
<td>3.37</td>
<td>.000</td>
<td>2.31</td>
<td>3.22</td>
<td>.000</td>
</tr>
<tr>
<td>Increases Research Activity</td>
<td>1.57</td>
<td>2.82</td>
<td>.000</td>
<td>1.77</td>
<td>2.60</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 4.4.1 The means and significance levels of the satisfied/dissatisfied group and the new system/not a new system group regarding publication statements.

Furthermore, a t-test was carried out on statements regarding publication in order to detect differences between those groups claiming to be either satisfied or dissatisfied with the system, and those wanting or not wanting a new system for evaluation. A significant difference was found in all instances, with significance levels ranging from $p < .000$ to $p > .049$. The means and levels of significance for those groups can be viewed in table 4.4.1 above.

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>Dissatisfied</th>
<th>Sig. Level</th>
<th>New System</th>
<th>Not a New System</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increases Quality of Publications</td>
<td>2.38</td>
<td>3.85</td>
<td>.000</td>
<td>2.74</td>
<td>3.54</td>
<td>.004</td>
</tr>
<tr>
<td>Provides a Correct Image of Quality of Dep. Pub. Outp.</td>
<td>2.65</td>
<td>3.97</td>
<td>.000</td>
<td>2.82</td>
<td>3.82</td>
<td>.000</td>
</tr>
<tr>
<td>Should Use Citations in a Purposeful Manner</td>
<td>2.67</td>
<td>2.87</td>
<td>.236</td>
<td>2.46</td>
<td>3.01</td>
<td>.038</td>
</tr>
</tbody>
</table>

Table 4.4.2 The means and significance levels of the satisfied/dissatisfied group and the new system/not a new system group regarding quality and citation statements.
Again, a t-test was performed to detect differences between those groups mentioned above in regard to statements on quality and citation use. A significant difference was found in all instances except one, as those either satisfied or dissatisfied with the system were equally for or against a purposeful use of citations within the evaluation system. The means and levels of significance for all groups in question can be viewed in table 4.4.2 above.

4.5 Collaboration

The next section of questions centred on the possible effects of the PBES on collaboration. Respondents were asked if they experienced the system as a factor in decreasing collaboration within the university itself, with other domestic universities, foreign universities and the Icelandic economy. Non-responses to these two questions ranged from 10-11%.

Asking respondents if they thought the system had a decreasing effect on collaboration within the institution, resulted in less than a quarter, or 22%, somewhat or strongly agreeing. On the other hand, 42% disagreed, either strongly or somewhat, with that statement. As a result, a quarter of the participants opted for the neither nor option.

When asking about possible effect on collaboration outside the university, 18% perceived the system to be decreasing collaboration between the UoI and other domestic universities. However, a much larger proportion of respondents, or 42% believed so not to be the case. Participants that neither agreed nor disagreed with the statement constituted the largest single category, or 30%.
Answers to the next statement regarding decreased collaboration with foreign universities had an even less support than the one pertaining to domestic collaboration. Only 14% of respondents believed the system to have a hand in less collaboration with foreign universities, whilst a majority of the survey participants (54%) either somewhat or strongly disagreed with that particular statement.

Furthermore, respondents were asked if they agreed or disagreed with the notion of decreased collaboration with the Icelandic economy. About a sixth of the respondents (16%) agreed with that statement to a somewhat or a strong degree, but a much larger proportion of 42%, did not. A third of the staff members neither agreed nor disagreed with the statement (31%).

Probed in an open ended manner about whether the PBES influenced collaboration in any other way than decreasing it, 33 respondents out of 74 believed so not to be the case.
collaboration. Most of these were further elaborations on decreased collaboration, such as the fractionalised counting of multi-author publications and decreased collaboration with the economy.

In addition, the PBES was seen as a factor in the choice of collaboration partners, with researchers rather opting for collaboration with those more prone to dish out results quickly for a quicker collection of research points. Also, the system was perceived as possibly encouraging university researchers to collaborate with students on their assignments, but simultaneously discourage them to recruit students for collaboration in their own research. Lastly, those reporting positive views were four. Those remarks centred on the system reinforcing collaboration between departments and increasing research - and hence increasing collaboration.

Moreover, respondents, who believed it to be the system’s role to encourage collaboration, were asked how the system could fulfil that role. Respondents mentioned reward for collaboration, such as interdisciplinary research and collaboration contracts, less punishment for multi-author articles via the fractionalised counting, and more points for articles in journals which emphasise multi-authorship. In specific relation to collaboration with the Icelandic economy, a respondent mentioned that those staff members in half positions or less should be incorporated into the system, as these parties usually have a stronger relation to the economy. This would then result in an increased collaboration with the economy on behalf of those staff members, as they would be prompted to do research because of the PBES.

When looking for difference between groups with a one-way ANOVA test a significant effect of the position of respondents on the view on decreasing collaboration in regards to Icelandic economy was reported, $F(4, 172) = 2.453$, $p < .048$. In order to discover
which groups were affected post hoc test were performed – Tukey HSD, Ryan-Einot-Gabriel-Welsch Range (REGWQ) and Games-Howell. The Tukey test showed no significant difference between groups, but the Games-Howell test revealed a significant effect between lectors and adjuncts (p < .044). Furthermore, the Tukey and REGWQ displayed a 3.00 mean for lectors and a 4.14 mean for adjuncts.

An effect of age was also found on the same statement of collaboration where $F(3, 172) = 3.840, (p < .011)$. The Tukey test then revealed a significant effect between the age groups 41-50 years and 60 years and older (p < .010). In addition, the Games-Howell test reported an effect between 31-40 years group and the 60 years + group (p < .034). Also, means of these groups varied slightly, with Tukey and REGWQ revealing a 3.13 mean for respondents age 41-50, a 3.20 for the 31-40 group, and finally a 3.91 mean for 60 years or older.

The age was also a factor in questions regarding collaboration with other Icelandic universities. The ANOVA reported a significant effect of the age of respondents on the view of decreasing collaboration with other Icelandic universities, $F(3, 173) = 3.599, (p < .015)$. The assumption of homogeneity was, however, violated; therefore, the Brown-Forsythe and Welch $F$-ratio were reported. The significant effect of age on the abovementioned statement was still present with $F(3, 78,832) = 3.835, (p < .013)$ (Welch) and $F(3, 150,322) = 3.911, (p < .010)$ (Brown-Forsythe).
As a result, the appropriate post hoc test of Games-Howell was performed, which revealed a significant difference between the age group of 41-50 years and 60 years and older (p < .009). In addition, the means for these two groups were reported by Tukey and REGWQ; the 41-50 years with a mean of 3.02 and the 60+ group with 3.79.

<table>
<thead>
<tr>
<th>Decreases Collaboration...</th>
<th>Satisfied</th>
<th>Dissatisfied</th>
<th>New System</th>
<th>Not a New System</th>
</tr>
</thead>
<tbody>
<tr>
<td>...within the UoI</td>
<td>4.01</td>
<td>2.90</td>
<td>4.03</td>
<td>2.93</td>
</tr>
<tr>
<td>...with Other Domestic Universities</td>
<td>4.02</td>
<td>3.03</td>
<td>4.15</td>
<td>3.07</td>
</tr>
<tr>
<td>...with Foreign Universities</td>
<td>4.21</td>
<td>3.28</td>
<td>4.26</td>
<td>3.40</td>
</tr>
<tr>
<td>....the Icelandic Economy</td>
<td>3.97</td>
<td>3.13</td>
<td>4.26</td>
<td>3.01</td>
</tr>
</tbody>
</table>

Table 4.5.1 The means of the satisfied/dissatisfied group and the new system/not a new system group regarding collaboration statements.

Furthermore, a t-test was performed, comparing those claiming to be either somewhat or very satisfied with the system as one group, and another group consisting of those either somewhat or very unsatisfied. A significant difference was found in relation to all statements regarding collaboration (p < .000). The same significance in group differences was found when those who either very or somewhat agreed with the implementation of a new system as one group, and those who very or somewhat disagreed as another (p < .000). The means for these groups on the appropriate statements can be viewed in table 4.5.1 above.

**Equality**

Next respondents were presented with statements regarding equality or inequality within the PBES. Firstly, staff members were asked if the system treated all disciplines equally. A clear majority of 60% disagreed, and even 34% did so strongly. When provided with the opposite statement, about inequality being apparent within the system, the same trend followed; 56% agreed, whereas only 15% disagreed. Lastly, respondents were asked if the system presented an inequality against specific disciplines. A majority of 52% agreed with the statement, and a third did so strongly. The non-responses for these questions ranged from 19-20%.
Chart 4.6.1 Respondents’ views on equality/inequality within the PBES.

When asked about which disciplines were affected by the system’s inherent inequality, the answers reflected the whole spectrum of disciplines within the university. The disciplines mentioned most often were those sharing a multi-author tradition in publications, such as medicine, biomedicine and physical sciences. Humanities were mentioned by eight respondents and experimental science by six. There were also six respondents who believed disciplines focusing on specific Icelandic subjects were discriminated against and five reported disciplines which publish few, but large, pieces of work. In addition, five respondents stated disciplines primarily publishing internationally, another five mentioned social sciences, as well as mathematics. Finally, disciplines where ISI publications are not seen as reflecting quality publications were reported by three respondents.
Respondents were then asked to identify the portrayal of inequality within the PBES. Here are some of the responses reported by the survey’s participants:

- Rewarding points for “worthless” publications.
- Not extensive enough, e.g. not rewarded for work outside the university directly related to research field – real life advisory roles or panels/committees.
- Icelandic journals are underestimated.
- Research activity different in each discipline, as a result opportunity to collect points is inherently unequal.
- Overestimation of ISI publications – might direct social sciences and humanities towards publications in these channels, although those are not the channels reflecting quality in these specific academic fields.
• Laborious and time-consuming research underestimated.
• Impossible to compare all these different disciplines, as they all have different patterns of publications. This difference should be considered.
• Researchers being registered as authors of a multi-author article, despite having not participated in the writing of the article.
• Underestimation of international publications.
• Fractionalised counting for multi-author articles not fair.
• Publications in Iceland in Icelandic journals overestimated.
• The role of lead author in multi-author articles underestimated.
• ISI has an inherent language bias.
• JIF and citations should be considered when evaluating articles.
• The work concerning collection of funds and operation of a laboratory is not taken into account.
• Not a tradition to refer to new pieces of research, as a result an inequality becomes apparent when citations are measured.
• Citations are no measurement of quality.

A one-way ANOVA was also performed, which revealed a significant effect of position on the view regarding inequality within the PBES, $F(4, 156) = 2.541 (p < .042)$. The Games-Howell post hoc test showed a difference between lectors and adjuncts with a sig. level of $p < .044$. Furthermore, means for these two groups were stated by Tukey and REGWQ to be 1.86 for lectors and 3.29 for adjuncts.

A significant effect of position was also found on the statement of inequality towards specific disciplines, $F(4, 155) = 2.866 (p < .025)$. The Games-Howell post hoc test reported a
significant difference between the groups of docents and adjuncts ($p < .042$), the means of those groups, as stated by Tukey and REGWQ were 1.90 for docents and 3.29 for adjuncts.

The same significant regarding inequality towards specific disciplines was reported when viewing groups defined by research obligation, $F(4, 153) = 2.528$ ($p < .043$). The Games-Howell test then revealed a significant difference between those with none research obligation and the 80% research obligation group ($p < .016$), between the 40% group and the 80% group ($p < .000$), and lastly those who reported other percentages of research obligation and the 80% group ($p < .000$). Furthermore, the means of those groups were reported by Tukey and REGWQ; the 80% group with a mean of 1.00, the 40% group 2.16, those with no research obligation 3.29 and those reporting other percentages 2.63.

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>Dissatisfied</th>
<th>Sig. Level</th>
<th>New System</th>
<th>Not a New System</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treats All Disciplines Equally</td>
<td>3.53</td>
<td>4.57</td>
<td>.000</td>
<td>3.56</td>
<td>4.31</td>
<td>.002</td>
</tr>
<tr>
<td>Inequality Is Apparent within the System</td>
<td>2.66</td>
<td>1.93</td>
<td>.002</td>
<td>2.87</td>
<td>1.98</td>
<td>.001</td>
</tr>
<tr>
<td>Inequality in Relation to Specific Disciplines</td>
<td>2.62</td>
<td>1.97</td>
<td>.007</td>
<td>2.92</td>
<td>1.96</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 4.6.1 The means and significance levels of the satisfied/dissatisfied group and the new system/not a new system group regarding equality statements.

In addition, a t-test was carried out in order to detect difference between those either satisfied or dissatisfied with the system regarding their view on equality or the lack thereof within the system. The same analysis was applied to those groups wanting a new system and those disagreeing with that change. A significant difference was found in relation to all statements ($p < .000$ to $p < .007$). The means for these groups for each of the statements can be viewed in table 4.6.1 above.
4.7 Research Behaviour

In order to grasp the respondent’s experience of the system in relation to research behaviour, statements touching upon possible effects of the system in those matters were introduced. About a fifth of the staff members (22%) believed the system to have a direct impact on their research behaviour, 19% neither agreed nor disagreed, but 39% disagreed. The next statement, which stated that because of the system the staff member places more emphasis on research, was also agreed upon by a roughly fifth of the respondents (22%). A third disagreed, and about 26% neither agreed nor disagreed.

The third question asked respondents if they would do less research were it not for the system. About 39% agreed with that statement, whereas 23% shared the opposite opinion. Lastly, staff members were asked about their risk in research in relation to the system. It is agreed upon by 32% of respondents that they take less risk in research because of the PBES. There were, however, 27% that neither agreed nor disagreed with the statement, and 20% believed so not to be the case. Non-responses in this section amounted to 22%.

Chart 4.7.1 Respondents’ views on the PBES’s possible impact on their research behaviour.
Furthermore, a one-way ANOVA test was conducted to determine different views between demographic groups. Firstly, a significant effect was found of the position of respondents within the university on their view regarding less research activity were it not for the PBES, \( F(4, 151) = 3.255 \) (p < .014). The Tukey post hoc test revealed a significant difference between adjuncts and those reporting other positions (p < .007), the same difference was found using Games-Howell, but to a less degree (p < .032). The means of these groups were then reported by Tukey and REGWQ: adjuncts with a mean of 3.86, whilst those with other positions had a mean of 1.83.

In addition, a significant effect was found regarding the age of respondents on their view on less risk in research because of the PBES, \( F(3, 150) = 3.687 \) (p < .013). Both post hoc tests, Tukey and Games-Howell, revealed differences between three different age groups. On the one hand between those aged 31-40 and 41-50 (p < .022 and p < .039), and on the other between 31-40 and 51-60 years of age (p < .028 and p < .050). Lastly, Tukey and REGWQ reported means for the three groups in question; the 31-40 group had a 3.43, the ones aged 51-60 had a mean of 2.52 and the 41-50 group 2.46.

Moreover, the one-way ANOVA revealed a significant effect of academic schools on the respondents belief that the PBES influenced their research behaviour, \( F(3, 147) = 4.105 \), p < .008. Both Tukey and Games-Howell post hoc tests indicated a difference between social sciences and health sciences (p < .014 and p < .013). Furthermore, Games-Howell also reported a significant difference between humanities and health sciences (p < .033). Tukey and REGWQ also disclosed the means for these groups, health sciences with a 2.89 mean and social sciences, as well as humanities, a 3.63 mean.

Lastly, views regarding the respondents added emphasis on research due to the PBES were found to be significantly affected by academic schools, \( F(3, 148) = 5.241 \), p < .002. The
post hoc tests of Tukey and Games-Howell revealed the difference to be between social sciences and health sciences (p < .021 and p < .012) on the one hand, and humanities and health sciences on the other (p < .006 and p < .005). Finally, Tukey and REGWQ reported the means of the groups in question, showing a 2.72 mean for health sciences, 3.47 for social sciences and 3.74 for humanities.

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>Dissatisfied</th>
<th>Sig. Level</th>
<th>New System</th>
<th>Not a New System</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Because of the System I Put More Emphasis on Research</td>
<td>3.51</td>
<td>2.89</td>
<td>.006</td>
<td>3.51</td>
<td>2.79</td>
<td>.002</td>
</tr>
<tr>
<td>Were It not for the System I Would Do Less Research</td>
<td>3.15</td>
<td>2.42</td>
<td>.002</td>
<td>3.44</td>
<td>2.30</td>
<td>.000</td>
</tr>
<tr>
<td>I Take Less Risk</td>
<td>2.90</td>
<td>2.74</td>
<td>.518</td>
<td>2.72</td>
<td>2.74</td>
<td>.938</td>
</tr>
</tbody>
</table>

Table 4.7.1 The means and significance levels of the satisfied/dissatisfied group and the new system/not a new system group regarding research behaviour statements.

Finally, a t-test was performed in order to detect differences between those either satisfied or dissatisfied with the system regarding their answers on research behaviour. A significant difference between groups was detected in relation to all statements, apart from the one pertaining to risk taking. Furthermore, the same analysis was applied to those longing for a new system, as well as those with the opposite view. That analysis resulted in only two statements having a significant difference when group answers, those regarding the system’s ability to encourage emphasis and the one pertaining to research behaviour. The means for these groups for each of the statements can be viewed in table 4.7.1 above.

### 4.8 A New System

The last section of the survey centred on the idea of introducing a new system for research evaluation. Those who agreed constituted a fifth of the staff members, 7% agreeing strongly with the suggestion. A considerable larger group of 42% disagreed, and the largest
single group of 26% strongly disagreed. Around 16% neither agreed nor disagreed with the idea of introducing a new system.

![Bar chart showing respondents' views on the implementation of a new evaluation system. N=154.](chart)

**Chart 4.8.1 Respondents' views on the implementation of a new evaluation system. N=154.**

Those respondents who believed that the UoI should use another system were the probed for ideas of a new system in an open ended manner. The responses could roughly be categorised into three different categories. The first one centred on the form of the evaluation, and the fact that the system could never successfully serve all disciplines equally. In this respect an idea of a comparative evaluation between departments of different universities was put forth. In that sense each discipline could be evaluated individually in comparison to the same disciplines within foreign universities. As a result, an evaluation of a single discipline would be carried out with a reference to international standards.

Others believed that the evaluation should be based on peer-review, for example in the form of committees or groups for either staff or research output evaluation. This was also mentioned in relation to a lack of knowledge, as foreign experts would serve better in certain disciplines, where little or no knowledge was present in the Icelandic science community. The
idea of just demanding a minimum research activity and stop at that was also mentioned, as well as less frequent evaluation, perhaps every five years or so. Additionally, some believed that the evaluation models of neighbouring countries, for example in the Nordic countries or Europe, should be consulted when establishing a new system.

Another group believed that a new system should change its emphasis and form of measurement. Respondents mentioned more emphasis on international research activity, as well as more emphasis on quality rather than quantity. Increased weight should be placed on ISI publications in highly prestigious journals, and peer-review demanded for all journals evaluated within the system. In terms of differences of measurement respondents reported a more citation-based system, and a system that evaluates funding awarded to researchers, with an extra weight given to foreign funding. The quality of different journals and book publishers should also be more actively taken into account. In addition, an added coverage and flexibility was mentioned, this centred on the different form of research output being evaluated, not simply written work. The outputs reported by respondents were, for example, video series or websites, teaching material, committee or workgroup participation, an outreach to the public and the implementation of new knowledge.

The answers of the last group were categorised as relating to the system’s connection to salaries and bonuses of staff. Firstly, the idea of an evaluation system being linked to salaries was protested. Secondly, respondents mentioned the decrease or elimination of the bonus pot, with the system only affecting the staff’s placement within the basic salary steps. An idea of the system rewarding departments or institutions instead of individuals was also proposed, as well as the idea of staff members being able to opt out of teaching obligation in return for a higher research obligation should they be successful in research.
Lastly, when conducting a t-test, in order to determine if those either satisfied or dissatisfied with the system were more likely to want a new system for evaluation, a significant difference between groups was detected (p < .000). The mean for the satisfied group amounted to 2.81, and the mean of dissatisfied respondents to 4.23. On a final note it must be noted that no difference between demographic groups were detected according to the one-way ANOVA test when viewing the answers to the new system question. In addition, no difference was found between the answers of male and female respondents in any of the question in the questionnaire according to an independent sample t-test.

4.9 Summary

This chapter of the dissertation outlined the results of a questionnaire presented to the personnel at the UoI. The population consisted of 776 staff members and the response rate amounted to 26%. Demographic questions revealed the academic positions and schools, research obligation, gender and age of the participants, which was then used as means of further analysis of responses made by each demographic group.

The answers to the questions regarding the respondents’ views on the system yielded some interesting results, as it became apparent that almost the equal percentage of respondents was satisfied with the current system, as was dissatisfied. When summarising the results, the descriptive statistics showed that staff members believed the system to increase publications, but not quality. In addition, a more active use of citations was agreed upon. In terms of collaboration, the respondents did not see the system as having a decreasing effect, but when moving towards questions of equality most staff members believed the system to be inherently unfair.

The answers to the questions regarding impact on research behaviour were not as homogeneous, as respondents disagreed with the notion of the system impacting their research
behaviour and emphasis on research. They did, however, agree when presented with statements regarding the system’s ability to encourage research activity, as well as its tendency to reduce risk in research. Lastly, the majority of the personnel was not interested in the implementation of a new system.

Last, but not least, the responses were analysed as a way to compare the difference of opinion between the demographic groups. The tests revealed a difference in means in quite a few the statements, for example, by position, age, academic school and research obligation. No significant effect was, however, found between gender groups. Furthermore, the groups of those either satisfied or dissatisfied, as well as those longing for new system or not, were used for a similar analysis, which also reported significant differences in most instances. Those results, as well as other presented within this chapter and last, will be further discussed in the next chapter.
Chapter 5 – Discussion

This last main chapter of the dissertation centres on discussing the results obtained from both the bibliometric analysis, as well as the questionnaire. Firstly, the bibliometric analysis will be discussed; the anomalies or interesting aspects within the comparative analysis, as well as the status of the UoI, especially in relation to the PBES. Secondly, the intriguing outcomes of the questionnaire will be debated and compared to the bibliometric analysis findings. Thirdly, the research outcomes will be applied to the programme impact theory developed in the methodology chapter. Lastly, final observations and recommendations will be made, followed by a few methodological reflections.

5.1 Publication Analysis

When viewing the outcome of the comparative publication analysis, the immense growth in publications during the 1994-2002 period at UoT is what stands out. Attempting to determine the reason for this outstanding upward development at Tromsø, the number of publications in each of the Noria-Net categories was examined. This was done to see if one could credit the growth to the university’s specific emphasis on disciplines known for rapid and frequent publications, such as medicine. Comparing the categorical profile of the top five publication categories did, however, reveal that Tromsø’s research profile was not significantly different from the other universities in question, and did, therefore, not provide and explanation for the phenomenon.

Explanation was then sought by viewing the index growth numbers for the institution. UoT has a sum growth of 1699% for the period, which is roughly 600% above the average (1105%) and the highest value of the six universities. Furthermore, Tromsø’s strength seems to lie in the fact that out of the 14 categories, it experiences a growth above the average, ranging from 8% to 193%, in 11 categories. The majority of the institutions are far behind Tromsø in its categorical growth range; Bergen has a growth above the average in two
categories, Lund in four, Uppsala in three and Copenhagen in five. The UoI comes closest to Tromsø with eight categories presenting publication growth exceeding the average.

In relation to the status of the UoI, the institution seems to follow a similar growth trend as the other universities. It is difficult to determine if the rise in 1998 is significant or not as the total publication numbers are relatively small and, therefore, prone to fluctuate. The same is true of the decrease in the following year. One could, however, speculate if these changes can be contributed to the PBES, as Gunnar Sivertsen introduces a similar trend in relation to attitude towards the Norwegian Model (Sivertsen, 2009). Sivertsen presents a negative curve in attitude the year after implementation, that is, as changes become visible in budgets (2009). In this instance, negative reactions could be interpreted as resulting in less research activity, although such ideas are simply speculative and perhaps farfetched.

As reported before, the UoI is doing relatively well in relation to growth, as it has values exceeding the average in 8 categories out of 14 with a total growth percentage of 1587, that is, 482% above the average. Although, as has been stated before, it is difficult to generalise with such a small dataset, these numbers clearly show that considerable growth in publications has taken place within most disciplines between 1994 and 2002. Those categories lagging behind are those of multidisciplinary science (-191%), mathematics (-128%), social sciences (-41%), physics (-37%), biomedicine (-32%) and ICT (-23%).

The reason for the decrease in publication numbers of these specific disciplines can partially be explained by their publication patterns; as some might tend to have a “slower” pattern and the period under investigation is quite short. In addition, the platform used to conduct the analysis is also too centred on the “hard sciences”, not giving a full or fair portrayal of categories such as social sciences and ICT, which tend to present their findings in
other forms not covered by WoS (although this is possibly changing with the inclusion of conference proceedings in 2008 (Thompson Reuters, 2011).

Determining growth in relation to the PBES was also attempted by looking at growth four years prior to the implementation of the system and then four years afterwards. The results seem to indicate that the PBES is not affecting publication behaviour as intended. Only five categories show numbers above the average when it comes to growth and only two of them in any substantial amount – biomedicine 16% above the average and social sciences 18%. As a result, these numbers indicate that the system is not contributing towards larger amounts of publications in general.

Contrastingly, the outcome of the questionnaire shows that staff members view the system as a stimulating component in research output. This same attitude was evident when Sigfúsdóttir et al. conducted interviews with staff members as a part of their evaluation (2005). The difference between these two results could be contributed to two things. Firstly, comparing these two results could in some way provide a distorted picture, as the bibliometric analysis focuses on a period perhaps too far away from the current situation, as one would expect most respondents to be commenting on their views and experiences in current time. Additionally, some staff members might not have been employed within the time frame the bibliometric analysis is set. As a result, should the bibliometric analysis have a larger time frame, one closer to the current time, it might provide a different picture of this specific aspect of the system.

Secondly, the reason could also be grounded in the platform used for the bibliometric analysis. A rise in publications could in fact have taken place, but not within publication channels listed in the WoS. As a result, the outcome would indicate a rise in lower level publications, at least for those disciplines whose most prestigious publication channels are
presented in WoS. This view was also aired in open-ended responses within the questionnaire. Consequently, one starts to wonder if the PBES is steering the UoI towards another Australian RQ. This will be discussed further in the citation analysis section, where the citation results will be introduced.

As presented in the results section the answers to the publication questions varied slightly in relation to age, academic schools and position, but perhaps the most intriguing difference is the one between groups of satisfied/dissatisfied system users, as well as the new system/not a new system group. Interestingly, answers did not reveal a great difference as the satisfied group strongly agreed with the statement of the system’s ability to increase publications, and the opposite group somewhat agreed. In addition, the satisfied respondents somewhat agreed with the notion of the system providing a correct image of a department’s publications, as well as the PBES’s ability to increase research activity. The dissatisfied respondents neither agreed nor disagreed with those statements.

When viewing the groups in relation to their longing for a new system, those in favour of a new system do in fact believe the current system to have the effect of increasing publications, providing a correct image of a department’s output, as well as increasing research activity. Surprisingly, those content with using the current PBES are not as sure of the system’s ability to fulfil the goals incorporated in the aforementioned statements. Nevertheless, the group somewhat agrees with the statement regarding publication increase and the mean for the other statements represents the neither nor option.

What this part of the study reveals is that those dissatisfied with the system are not so because of its ability to encourage and motivate researchers to publish and do research. The same applies to those respondents who want a new system. Their dissatisfaction must,
therefore, be rooted within other aspects of the PBES, which hopefully will be uncovered during the next sections.

5.2 Collaboration

Viewing the numbers for collaboration revealed some interesting points of discussion, especially in relation to the UoI. The collaboration analysis indicates that the UoI rises slightly above the average when the proportion of articles containing collaboration is examined. The reason for this development could simply rest on Iceland’s small size. The fact that the UoI is by far the largest university in Iceland might make national collaboration not as accessible or attractive as would possibly be the case with the other Nordic universities. This would be especially true for those disciplines only taught at the UoI, with no possible collaboration partners than those abroad, except in cases of interdisciplinary research. This would, for example, apply to medicine and the majority of disciplines within the humanities.

The analysis also tried to establish changes in growth before and after the implementation of the PBES. The growth at the UoI amounted to 47%, which was 8% above the average of 39%. Although 8% can hardly be regarded as a great significant change between the two periods, the numbers obtained in the bibliometric analysis reflect the views of the staff members, who, by large, did not agree when asked if the system decreased collaboration.

The answers varied slightly when comparing different groups in relation to decreasing collaboration with the Icelandic economy as lectors neither agreed nor disagreed, whereas adjuncts tended to somewhat disagree to the statement. In addition, age was also reported as a significant factor, with those aged 31-50 neither agreeing nor disagreeing, but the older group of 60+ leaned towards somewhat disagreeing. This same tendency was reported in the
question regarding decreasing collaboration with other Icelandic universities, that is, the age group 41-50 tended to neither agree nor disagree, whereas the 60+ group somewhat disagreed.

Additionally, a significant effect was found in relation to the respondents’ satisfaction with the PBES and their interest in setting up a new system. Unsurprisingly, those satisfied with the system somewhat disagreed with the statements of decreased collaboration. Those dissatisfied did neither agree nor disagree. Similarly, those who did not see the system as decreasing collaboration were not as enthusiastic about setting up a new system, whereas the opposite group tended to neither agree nor disagree.

This correlation between groups is in some ways easily explained, as one would not be surprised that those dissatisfied, and in want for a new system, would not be as positive in their review on the system’s possible negative impact. The response difference between staff status and age is, however, more difficult to explain, as it cannot be credited to any dissimilarities of treatment within the system. In terms of research obligation, those in full-time positions as adjuncts generally have a research obligation of 31%, whereas the research proportion for lectors is 43% (UoI, 2006b). One could try to argue that the time devoted to research and the more hands-on experience with the system has a negative impact, but that hypothesis does not hold, as that would presumable mean that professors would agree even more, when they in fact have a mean of 3,60 (somewhat agree). It must, therefore, be concluded that this difference is probably due to characteristics of too small a sample size to give enough representation to each age group.

5.3 Citation Analysis

The comparative bibliometric analysis of citation impact reveals a rather interesting development for the UoI. According to the growth numbers obtained from the analysis, the UoI is lagging considerably behind, which is a completely different trend from the one
revealed by the publication analysis, where the institution had a combined growth exceeding the average by 482%.

Nevertheless, this does not seem to be a clear by-product of the PBES, although its implementation has not succeeded in turning this development around. It is evident that the citation impact is already on the downward spiral in 1997, but does, however, reach the average in 1999, only to decrease again and stay below the average until the end of the investigative period. This is made even clearer when viewing the Crown Indicator average before and after implementation, where the UoI is the only institution that has a higher average in the before period.

These results indicate that the system is not working as expected in regards to citation impact. The PBES is constructed in a similar fashion to the Norwegian Model, that is, tries to encourage publications in highly prestigious publication channels by offering more research points for material published therein. The root for this malfunction might be found in how many points are awarded for top notch publications in the highly esteemed journals or publication houses.

As an example, one can look at the list of journals, where a researcher receives 20 research points for an article either in an ISI journal, with a journal impact factor high enough to be placed amongst the top 10% in its category, or in an A-category ERIH journal. For only five points less an article in any other ISI journal will suffice, along with any first class Icelandic journal or B-category ERIH journal. The system then awards ten points for second class Icelandic journals and C-category ERIH ones. Lastly, one can obtain five points for third class Icelandic journals.

As a result, the incentive to conduct an extensive and time consuming research, which would be accepted into a top ISI or ERIH journal, is to some extent lost. The same researcher
could just as well publish two articles in a second class Icelandic journal in order to obtain the same amount of research points. This is not saying that the work is any less extensive or time consuming when published in a second class Icelandic journal, but it is, however, safe to say that the competition for a slot is less in a journal within a research society consisting of roughly 2300 researchers⁵ and a speech community of approximately 300 thousand.

The fact that the system is based on personal gain, instead of departmental or institutional, might also play a role in the aforementioned development. As a result researchers might be more prone to produce an abundance of lower class publications for a quick collection of points. Of course this could also occur with an institutional or departmental funding system, as evident from the Australian Research Quantum, but the fact that publications can affect a person’s livelihood in such a direct manner, and that the system was partly thought of as a method to increase the salaries of professors when originally launched, reinforces this argument.

Currently, the PBES does not actively use citations when evaluating or awarding research points. One of the interviewees did, however, add that citations were taken into account when a person was up for a promotion. The survey, therefore, asked respondents if they thought that citations should be used in a more purposeful manner within the system. Almost 40% believed the system should use citations, whereas 26% did not agree.

The answers differed slightly when looking at academic schools, where the health sciences were somewhat more positive in respect to citation use than the other academic schools. The reason might be grounded on the fact that of all the subject categories examined in the citation analysis, clinical medicine at the UoI had by far the largest amount of citations each year, and, according to the recently released bibliometric report of the Icelandic Centre

---

⁵ Numbers from 2008 by Eurostat (Eurostat, 2008).
for Research, constitutes around third of all peer-reviewed publications in Iceland (2010). The use of citations would, therefore, presumably result in a greater gain by authors writing in this particular field.

A significant difference was also found between other groups, when asking about the quality of publications. In retrospect additional questions should have been added, which probed respondents on their view on the system’s relation to impact of publications in addition to quality. This would perhaps have been more appropriate in order to directly connect the statements to citations, as the relations between citations and quality are by no way concrete (see Moed, 2005).

Nevertheless, viewing the results from the quality statements, the most interesting outcome was perceived in groups of satisfied or dissatisfied with the PBES. The satisfied group somewhat agreed with the system’s ability to increase the quality of publications, as well as the fact that it gives a correct image of the quality of the material published by the departments. Unsurprisingly, the dissatisfied group somewhat disagreed with those statements.

5.4 Research points and Equality

In this part of the analysis, the focus was shifted from the comparative angle to a more specific view on the UoI. Firstly, information on the research points awarded to academic schools at the university was collected. These numbers presented two blocks; social sciences and engineering and natural sciences with roughly 27.400 research points each year, and then health sciences and humanities with approx. 24.800 points annually.

The reason for the difference could either be a mere difference in publication quantity, or an inherent inequality with its basis in publication format and the research points awarded there for. Arguing inherent inequality might, however, be a challenge, at least when basing
the argument on the publication format information. When viewing the publication format of the academic schools it becomes apparent that the “hard sciences” with similar publication patterns (health sciences and nature sciences), as well as the “soft sciences” of humanities and social sciences, each have their representation within each block.

When viewing the publication format of those two publication pattern groups for 2006 to 2008, it becomes evident that differences lie mostly in publication of book chapters and ISI articles. The academic schools differ quite extensively in this respect, with 1% to 8% of the “soft sciences” publishing in ISI publication channels, but 31% to 46% of the “hard sciences”. Additionally, the “soft sciences” have 24% to 31% of their publications in book chapters, but the “hard sciences” only 3% to 24%. Those differences in publication format do, however, not seem to affect the ranking of academic schools in terms of research points, as both book chapters, as well as ISI articles offer maximum research points of 20 points. This means that the inequality reported by staff members is not based in the format of the publications, but could, however, be rooted elsewhere.

Although these two arguments seem to rather strengthen the notion of a fair system, at least in relation to publication format, fairness is by no means the view or experience of the respondents of the questionnaire. The attitudes regarding the system’s equality were quite straightforward negative, that is, all statements had majority of respondents agreeing on the lack of equality. Interestingly, when conducting the one-way ANOVA in the questionnaire chapter, no significant difference was detected between the responses of any of the academic schools, which might indicate that all of them feel discriminated against, but perhaps differently and to a different degree.

Another interesting aspect became evident when the means for those either satisfied or dissatisfied with the system, as well as those wanting a new system or not, were compared to
the answers to the equality/inequality statements. The first assertion, stating that the PBES treats all disciplines equally, revealed that those satisfied with the system somewhat disagreed, and those dissatisfied disagreed strongly. Similarly, those wanting a new system somewhat disagreed to the statement, as did the ones preferring no change. Secondly, asked about inequality being apparent within the system, also in relation to specific disciplines, the satisfied group neither agreed nor disagreed, whereas the dissatisfied one somewhat agreed. The same answering pattern was found when looking at the new/not a new system groups, with those yearning for a new system neither agreeing nor disagreeing, and those wanting to keep the current PBES somewhat agreeing.

These results are quite astonishing, as one would imagine those satisfied with the system and those not wanting a new one, would answer differently. It seems rather strange that one would be satisfied with a system one also perceives to be unfair. The reason for these surprising replies could be twofold. Firstly, the satisfied group constitutes both those very or somewhat satisfied with the system, so that those “somewhat” satisfied might perhaps not be very satisfied due to the system’s perceived inequality. Secondly, the satisfaction could simply be rooted in something else than equality, and the preference for a new system rooted in other changes than those to do with fairness.

5.5 Impact on Research Behaviour

Further questions on research behaviour, which did not have an accompanying bibliometric analysis, were also presented to the respondents of the questionnaire. The answers to one of these questions have a rather contrasting tone to what has already been discussed above in relation to research behaviour. As had been established the majority of respondents believed the system to have a stimulating effect on publications, however, when asked whether the system had a direct impact on one’s research behaviour, only about fifth of the respondents believed so to be the case. The reason for these contrasting answers is
unclear, as in the next statement roughly 40% believed that were it not for the system, they would do less research.

One of the reasons might be that respondents do not perceive increased publications as a change in research behaviour. In addition, the discrepancy might lie in the wording of the questions, as the questions in the publication chapter are less personal and more general, whereas the statements on research behaviour were more individually directed. As a result, one might perceive the system as being stimulating publications in general, but not applying to oneself. Furthermore, and perhaps the most interesting aspect in need for further research, the general view might be a myth that respondents have heard around them and believed to be true, whereas not experiencing the same effect personally.

When viewing the statements in relation to whether respondents were satisfied with the system or if they wanted a new system yielded some interesting points. A significant effect was only found in relation to satisfaction with the system when respondents were asked about the system’s direct impact on research behaviour. Those satisfied somewhat disagreed with the statement, whilst the dissatisfied ones were quite neutral in this respect. Additionally, the trend is reversed when looking at the statement “Because of the system I put more emphasis on research”, where the group wanting a new system disagreed and the opposite group neither agreed nor disagreed. It must, however, be noted that a significant part of those dissatisfied with the system also make up a large part of the group which does not want a new system (see section 5.6). Finally, the last statement to have a significant difference in responses between groups was the one pertaining to less research were it not for the PBES. The satisfied respondents had a neutral position, as did the ones rooting for a new system. Surprisingly, those dissatisfied with the system somewhat agreed with the statement, as did the ones not desiring a new system.
These results present rather interesting aspects in relation to the overall view of those either satisfied or dissatisfied with the PBES. Firstly, those claiming to be satisfied rather disagree with the notions that the system has a direct impact on their research behaviour or is a determining factor in stimulating research. This comes across as being rather odd, seeing as those aspects relate to the central components or aims of the system. Of course it is not clear as to what parts of the system the group is particularly satisfied with, should it not be the system as a whole, and one could even speculate so that if the group sees the system as being so non-intrusive in relation to research behaviour, that this aspect might simply be what encourages this group’s satisfaction.

Another debatable point has to do with the opposite group of respondents, that is, the dissatisfied one. Contrastingly, this group, despite not being content, sees the system as affecting them, as were it not for the system this group would do less research. This is perhaps more understandable than the case above, as one could be dissatisfied with a number of things, but still admits to a certain effect the system has on research behaviour. This effect might even be a negative one. In addition, this reflects former responses in the publication chapter, where the majority, and hence also those dissatisfied, perceived the PBES as stimulating factor in research output.

Although no relation was found between different groups when asking about less risk behaviour in research because of the system, the responses were, nonetheless, interesting. In some way it can be said that those results reinforce preconceptions regarding a budget allocating research evaluation systems. In this instance 32% agreed when asked if they took less risk in research because of the PBES. This supports similar statements made by Aldo Geuna in relation to the economics of knowledge production (1999), and is by no means surprising as gambling with one’s livelihood is not a particular practical option.
5.6 The Overall Satisfaction and a New System?

The ultimate question of whether staff members were satisfied or not with the PBES system was the starting point for the survey, and revealed roughly equal groups of those satisfied or dissatisfied. When comparing means of groups, the health sciences were somewhat more dissatisfied than the humanities and social sciences, which occupy a neutral position on the matter. The reason for the dissatisfaction of this particular academic school might simply be rooted in the fractionalised counting of multi-author articles (see literature review), which was often aired in the survey as both a potential factor for decreased collaboration, as well as inherent inequality.

The most interesting aspect of the overall picture is that apart from roughly 40% being dissatisfied with the system, only 20% of all respondents agree with the idea of a new system. When looking at the two groups of either satisfied or dissatisfied respondents in relation to a new system, those satisfied tend to have a rather neutral position, whereas the dissatisfied group, and the group one would expect to be more interested in severe changes, tend to somewhat disagree.

Why respondents seem to be interested in keeping the current system, even those not happy with it, remains a mystery. It might simply be so that people feel safer attaching themselves to something they already know, even though it is not to their liking, as a new system might be even worse. In addition, the dissatisfied group might not see the system as a total failure, and would rather opt for a change in the current system, rather than moving onto a new system of evaluation.

5.7 The Theoretical Impact Model

Viewing the programme impact theory from the methodology chapter in the light of the information received from both the bibliometric analysis and the questionnaire, perhaps adds more questions than it answers. Firstly, by looking at the possible moderators
experienced by staff members at the university it becomes evident looking through the results of the collaboration and inequality analysis that those aspects can actually not be classified as moderators. The reason for such a conclusion is rooted in the fact that all those aspects mentioned by the staff members, as contributing to inequality, are in fact not outside factors, but inherent within the system itself. As a result, trying to determine how these perceived notions affect the system’s function is not possible, unless conducting another survey where respondents would be prompted to explain how exactly they saw these factors influencing their publication behaviour. That would, therefore, be a useful future step for a further analysis. In addition, the bibliometric analysis of publication formats and research points did not reveal an inherent inequality, which, however, does not say that it cannot be found elsewhere within the system.

Although the collaboration analysis did not reveal a significant change in collaboration after the implementation of the PBES, the UoI was slightly above the trend lines of the other institutions during the whole period of analysis. As discussed above the reason is most likely

![Image 5.7.1 A revised programme impact theory where moderators have been changed according to results obtained from the bibliometric and survey results.](image-url)
due to the size of the country and the subsequent need for foreign partnership. As a result, the collaboration cannot be directly perceived as being simply a moderator affecting the publications coming out of the system, but an essentially a part of the context the university and the system thrive in.

In terms of the system’s proclaimed goals of increasing publications in prestigious publication channels the pictorial impact theory can be either supported or disregarded, depending on which analysis is used for the portrayal. Looking at the views of the survey respondents the system does increase publications, which reinforces that part of the theory. If one, however, views the bibliometric numbers, the UoI has the same trend as other universities in publications, as well as no substantial growth after the system’s implementation. Those numbers would, therefore, indicate that publications come out whether the system is present or not.

Image 5.7.2 A revised pictorial representation of the programme impact theory. Here the relationship between the system and the motivational moderators has been eliminated, as results indicate that the system does not seem to increase publications in high impact publication channels.
The bibliometric citation analysis, however, revealed that the system was unsuccessful in encouraging publications in those publication channels perceived as being outstanding or prestigious. As explained above, citations are currently not being used in the research evaluation process, but the attempt to steer publications towards the “right” publication channels is based on the point system. This particular procedure does not seem to call forth the desired effect the system seeks, as citation growth lags behind the other Nordic universities, and no change is apparent after the implementation of the PBES. As a result, this eliminates, or at least weakens, the relationship between the PBES and the motivational mediator (see image 5.7.2 above).

5.8 Final Observations and Recommendations

On a final note the most interesting aspects revealed in the study are twofold. Firstly, the fact that the bibliometric analysis disclosed that despite publications being on the rise at the UoI, in a similar manner to the other Nordic institutions, the institution’s citation impact is not able to keep up with the increase of the other six universities. Although this development was not seen as direct effect of the PBES, the system's implementation has not aided in its reversal.

Nevertheless, the current system could be mended to reverse this trend. Firstly, it must be noted, that the difference in research points between those publications deemed “normal” and those seen as exceptional is too small. As discussed above, the fact that a researcher can easily collect as many research points for a publication in a prestigious publication channel as when publishing two or three in a second or third class national journals, does not encourage the extra work and effort needed for the exceptional piece of work. The PBES needs, therefore, to be changed as to make the gap between these two platforms wider, and thereby encouraging researchers to change their publication behaviour towards the higher level publication channels.
Another change which might indirectly serve as a mean to stimulate high impact research would be to move the bonus pot feature of the system to a higher level, that is, departmental or even institutional. The 12.5% bonus pot could then be incorporated into the basic salaries of the staff, and a part of the university’s research budget used for a departmental bonus pot. This change would remove the link from a researcher’s personal livelihood to a more collective approach, and would probably allow researchers to partake in a more laborious and risky assignments without risking their livelihood.

The most fascinating point observed, when moving on to the survey results, was, however, the fact that despite roughly half of the respondents being dissatisfied with the current system, only a fifth thinks that a new system should be introduced. In addition, the majority sees the system to be inherently unequal, which makes the results even more surprising. As was discussed above, these somewhat confounding results could be linked to the fact that people are indeed not exceptionally dissatisfied with the system on the whole, but rather have concerns regarding specific aspects of it.

The system itself will, perhaps, never be completely equal, at least, it will not be perceived so by all those affected by it. Regarding the difference in the speed and quantity of publication output between fields, one of the respondents suggested that departments be evaluated in comparison to similar departments in other universities. This could be a way to bypass the aforementioned difference, but could, however, also prove difficult to implement as finding comparable departments in all respects might be a challenge.

Additionally, those equality issues due to differences when choosing publication channels, could be bypassed by following the ideas of the Norwegian Model, where each discipline decides upon publication channels for level 1 and 2. This change would allow each
field to pick the most prestigious or sought after journals or publishers, instead of following the guidelines of ISI or ERIH.

5.9 Methodological Reflections

Having gone through both of the research strategies described in the methodology chapter, as well as contemplating the results within the present chapter, a few points were made to the structuring and implementation of each strategy. When processing the survey responses two distinct observations were made. Firstly, as discussed above, further questions regarding the impact of publications would have been optimal. Secondly, a seven response Likert scale would also have been a viable option. It might not have provided drastically different set of answers, but might serve as an interesting alternative option.

Attention was also brought to the conduction of the citation analysis within the bibliometric section. During the counting of citations, the number obtained by all articles, irrespective of their year of publication, were counted, which gave the articles 8 to 16 years to collect citations. A more consistent way might have been to allow each publication a five years citation window, so that all would have an equal amount of time to accumulate citations. Nevertheless, the same technique was applied to all universities, so no institutional bias was introduced.
Conclusion

The initial purpose of this dissertation was to attempt to determine if the current research evaluation system at the UoI was affecting the research behaviour of staff members. The aspects given particular attentions were those regarding publications, citations, collaboration and equality. Bibliometric analyses, as well as a questionnaire, were conducted for the purpose of seeking the answers to the aforementioned aim.

The bibliometric analyses revealed no obvious or direct links to changes in publication, citation or collaboration behaviour, which could be credited to the evaluation system. During the comparative analysis the UoI had a similar upward trend as the other Nordic universities in relation to publications; had steady collaboration numbers slightly exceeding the average; and only 5% growth in citations, compared to a 38% average. None of these trends could, however, be linked to the implementation of the PBES. Furthermore, publication format of those publications rewarded research points during a three year period was also examined, in order to detect any discrimination or inequality stemming there from. The format analysis did, however, not give credence to that idea.

The answers of the staff members of the UoI did, however, not always coincide with the results of the bibliometric analyses. For example, the personnel believed the evaluation system to have an increasing effect on publications, although the bibliometric numbers did not support that notion. Additionally, the format analysis was not able to detect inequality, despite numerous claims thereof by staff. It must, however, be noted that inequality could stem from other aspects within the system than publication format.

Staff did, however, agree with the bibliometric analysis of collaboration, as it did not see the system as a diminishing factor in both national, as well as, foreign collaboration. The bibliometric analysis did, however, only cover foreign collaboration. As noted above,
citations are not used actively within the system today, so asking staff about their view on citation usage in terms of increased impact was impossible. A 40% of respondents did, however, want citations to have a more active role in the current evaluation system.

Trying to determine the effects of a system, such as the PBES, is rather difficult, and in some ways almost impossible, as isolating the variable of the system is beyond the bounds of possibility. Firstly, there is an abundance of other factors which could contribute to the change in research behaviour of the scholars at the UoI. For example, one could name changes in policy within the university, changes in emphasis with a new rector, as well as policy changes and budget adjustments at the governmental level. All of these factors, and probably many more, would have to be isolated in order to give a precise picture of the effects of a single system. Nevertheless, the study conducted in this dissertation is able to contribute to the ongoing debate about research evaluation systems, and perhaps give an indicator of the course the UoI is on, and a reflection of staff experiences.

Further research would be feasible in the respect with the purpose of peeling back the layers and hopefully get as close to the core, in this respect the isolated system, as possible. Firstly, examining the status of the university using its own numbers of publications would be an excellent start. This would allow for all publications to be examined and, therefore, make a better judgement regarding various factors, such as; the credence of the stimulating publication effects of the system and level of national collaboration.

Secondly, conducting another questionnaire research a few years from now would also be interesting, as to see how or if the views of the staff have changed, as the current version of the system has only been in use for just over a year. Including questions regarding the personnel’s view on the system's ability to increase impact, as well as how staff sees the inherent inequality of the system affecting its research behaviour, would be optimal as a way
to fill into some of the gaps found in this study. Another survey would also be especially intriguing if any aspects of the system, or the system on the whole, would be altered. Furthermore, taking the evaluation towards an even larger level, examining simultaneously possible effects of policy, budget changes and changes in research emphasis would be ideal.

On a final note, it is clear that whichever course research at the UoI takes, it is certain that measurement of science and research activities will follow. What is perhaps the most important aspect of all is that whichever method used for measurement, the need for it to be transparent and monitored closely in an open fashion is crucial, as to reinforce the trust and satisfaction of those being evaluated by it. This would also mean further opportunities for bibliometricians to pursue and practise their evaluative techniques, hopefully resulting in a more fruitful bibliometric field in the near Icelandic future.
Bibliography


Meho, L. I. & Yang, K. (2007). Impact of Data Sources on Citation Counts and Rankings of LIS Faculty: Web of Science Versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58* (13), 2105-2125.


Hello,

You have been invited to participate in a survey on the effects of the performance-based evaluation system of the University of Iceland in relation to research. In this survey the academic staff of the University will be asked to answer questions relating to how the above mentioned system can potentially affect research aspects, such as cooperation, competition, publication and equality. It should take about 5-10 minutes to complete the survey.

Your participation in this study is completely voluntary. You are not obligated to answer all the questions, and can of course withdraw from the survey at any point. I would greatly appreciate your participation as your opinions are important. Please note that the questions are only aimed at your own experience and attitude, so no special knowledge of the system itself is required.

Your survey responses are of course anonymous and will only be used in relation to this specific dissertation. If you have any questions about the survey or the procedures, you can contact me, Eva Dögg Diego Þorkelsdóttir, via e-mail at k09evth@stud.iva.dk or evadth@gmail.com.

Thank you very much for your time and support. Please start with the survey now by clicking on the Continue button below.

Vinsamlegast veldu tungumál/Please choose a language
- Íslenska
- English

Which of the following best describes your position within the University of Iceland?
- Professor
- Docent
- Lector
- Adjunct
- Other

Which department do you, or the research institution you are affiliated with, belong to?
- Social and Human Sciences
Social Work
Economics
Law
Political Science
Business Administration
Nursing
Pharmaceutical Sciences
Medicine
Food Science and Nutrition
Psychology
Odontology
Foreign Languages, Literature and Linguistics
Theology and Religious Studies
Icelandic and Comparative Cultural Studies
History and Philosophy
Teacher Education
Educational Studies
Sport, Leisure Studies and Social Education
Industrial/Mechanical Engineering and Computer Science
Earth Sciences
Life and Environmental Sciences
Electrical and Computer Engineering
Physical Sciences
Civil and Environmental Engineering
Other

Please identify your research obligation
- None
- 40%
- 60%
- 80%
- Other

Please specify gender
- Female
Male

Please specify age
- 20 years and younger
- 21-30 years
- 31-40 years
- 41-50 years
- 51-60 years
- 61 years and older

How satisfied are you with the current performance-based evaluation system of research at the University of Iceland?

<table>
<thead>
<tr>
<th>Very satisfied</th>
<th>Somewhat satisfied</th>
<th>Neither satisfied nor unsatisfied</th>
<th>Somewhat unsatisfied</th>
<th>Very unsatisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Which of the following factors do you think is the most important in relation to research activity within the University of Iceland? (Please rank according to importance, 1 for most important etc.)

- Favourable research policy
- Increased funding for research
- Performance-based evaluation system
- Improved research facilities
- More PhD students
- More research staff
- Increased cooperation with other Icelandic universities
- Increased cooperation with foreign universities
- Increased funding available within competitive research funds within, as well as outside, the University of Iceland

Which of the following factors do you think is most important in relation to research quality within the University of Iceland? (Please rank according to importance, 1 for most important...
etc.)

Favourable research policy  
Increased funding for research  
Performance-based evaluation system  
Improved research facilities  
More PhD students  
More research staff  
Increased cooperation with other Icelandic universities  
Increased cooperation with foreign universities  
Increased funding available within competitive research funds within, as well as outside, the University of Iceland

<table>
<thead>
<tr>
<th>The performance-based evaluation system</th>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>decreases cooperation between researchers within the University of Iceland</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>decreases cooperation between researchers at the University of Iceland and researchers at other domestic universities</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>decreases cooperation between researchers at the University of Iceland and researchers at foreign universities</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>decreases cooperation between researchers at the University of Iceland and parties of the Icelandic economy</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>increases cooperation between researchers within the University of Iceland</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
increases cooperation between researchers at the University of Iceland and researchers at other domestic universities increases cooperation between researchers at the University of Iceland and researchers at foreign universities increases cooperation between researchers at the University of Iceland and parties of the Icelandic economy

Does the performance based evaluation system alter cooperation between researchers at the University of Iceland and other researchers (domestic or foreign) or parties of the Icelandic economy in any other way?

The performance-based evaluation system should, in a purposeful way,

encourage cooperation between researchers within the University of Iceland encourage cooperation between researchers at the University of Iceland and researchers at other domestic universities encourage cooperation between researchers at the University of Iceland and researchers at foreign universities encourage cooperation between researchers at the University of Iceland and parties of the Icelandic economy
If you believe the role of the performance-based evaluation system is to purposely encourage cooperation between researchers - how do you think the system should fulfil that role?

The performance-based evaluation system increases competition for funding from

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Iceland research funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other domestic research funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>foreign research funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The performance-based evaluation system should encourage

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>increased competition between researchers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>increased demands for funding from Icelandic competitive research funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>increased demands for funding from foreign competitive research funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you believe the role of the performance-based evaluation system is to encourage increased competition and increased demands for funding from competitive research funds - then how do you feel the system should fulfil that role?

The performance-based evaluation system

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>increases publication output</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>increases the quality of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
publications
gives a correct picture of the publication output of my department
gives a correct picture of the quality of my department’s publications
increases research activity
should use numbers of citations in a purposeful manner when evaluation takes place

### Equality within the system?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The performance-based evaluation system treats all disciplines equally</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality is apparent within the system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality is apparent in relation to specific disciplines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you believe the system to have an inherent inequality - which disciplines is that affecting?

<table>
<thead>
<tr>
<th>Disciplines</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

How is that inequality portrayed?

<table>
<thead>
<tr>
<th>Inequality</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

How strongly do you agree or disagree with the following statements?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Somewhat agree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The performance-based evaluation system has a direct impact on my</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
research behaviour
Because of the performance-based evaluation system I put more emphasis on research
Were it not for the performance-based evaluation system, I would do less research
Because of the performance-based evaluation system I take less risk in my research
The University of Iceland should use a different system of research evaluation than is presently used

If you believe the University of Iceland should use a different system to evaluate research - do you have an opinion on how such a system should be?

Any final comments on the performance-based evaluation system in relation to research, publication or anything else?
II. Appendix – Additional Visual Survey Results

Chart II.1 The number of respondents by academic department.
Chart II.2 Does the system increase competition for funding?

Chart II.3 Most important factors in relation to research activity at the UoI (%). This data includes answers from School of Education.
Chart II.4 How respondents rank the importance of the PBES as factor in research activity at the UoI. This data includes answers from School of Education.

Chart II.5 Most important factors in relation to quality of research at the UoI (%). This data includes answers from School of Education.
Chart II.6 How respondents rank the importance of the PBES as factor in the quality of research at the UoI. This data includes answers from School of Education.
III. Appendix – Evaluation System for Public Universities

Evaluation Criteria
The Evaluation System for Public Universities forms the basis for job evaluation of academic staff, i.e., persons hired in accordance with standard procedure (assessment of qualification) who are members of the Icelandic Association of State University Professors or other collective bargaining organisations party to the Evaluation System (including the Union of University Teachers, Teachers’ Association of the Iceland University of Education and Akureyri Union of University Teachers).

At the University of Iceland, publications and published materials not associated with the University of Iceland (Icel. Háskóli Íslands) are not evaluated. Evaluation covers research, teaching, administration and service work and other factors. Annual performance reporting takes the form of a report detailing activities during the previous year. At the University of Iceland, reporting also includes updating the teaching resume, an overview of collaboration with parties outside the University and an overview of work done outside the University. No points will be awarded before a satisfactory performance report has been submitted.

Research
Research evaluation is largely based on the publication outlet. Research appearing in a publication outlet that makes rigorous scholarly demands of its content is considered to have been already evaluated with regards to data solicitation, originality and contribution to advancement of knowledge. The peer-review process is a critical factor in publication. For peer review to be considered satisfactory, the material in question must be sent to a minimum of two reviewers. Peer review must be anonymous and professional. It must relate to content and be made by recognised experts in the field in question. Publication distribution, accessibility and impact are also evaluated.

Teaching
Teaching activity and quality are evaluated, including publication of teaching materials, innovation in teaching and supervision of graduate students. At the University of Iceland, points for teaching experience (B1) are not awarded unless the teaching resume has been updated (see above). This resume must be updated annually as part of the performance reporting process.

Administration, service, etc.
Administrative work is evaluated for points (cf. C), as is service work (cf. D). Points are given for specific administrative positions held within the university. Most positions evaluated under section C involve work under the aegis of the university as a whole or its various schools. Administrative work under the aegis of individual departments is not, in general, evaluated for points according to these rules. Public education work and services that build on an individual’s expertise are evaluated under section D.
A. Research

a) Division of points for multi-author material
For multi-author books (A2), book chapters (A3), academic articles (A4), conference proceedings (A5), editorial work on academic publications (A7), reports, reviews and translations (A8), teaching materials (A9) and innovations (A10), division of points is calculated as follows:
Two authors: 1.5 x points / 2
Three authors: 1.8 x points / 3
Four or more authors: 2.0 x points / number of authors

Academic staff may receive additional points for up to two multi-author works annually as outlined below:
1. If a staff member publishes four or fewer multi-author works, (s)he receives additional points for one work – that for which the person in question will receive the most points after calculation – amounting to half of the difference between points awarded and points prior to division between authors.
2. If a staff member publishes more than four multi-author works, (s)he receives additional points for two works, amounting for each to half the difference between points awarded and points prior to division (see above).

A staff member thus receives additional points for one work if (s)he publishes four or fewer multi-author works in a year and additional points for two works if (s)he publishes more than four. The number of additional points received increases as the number of authors rises.

Example:
A publication receives 20 points. Points are divided between six authors, each of whom is awarded 6.67 points (2*20/6=6.67). The difference between points awarded (6.67) and points prior to division (20) is 13.33 points. As per the rules outlined above, authors receive half of this difference (6.67 points) or a total of 13.34 points each. Should any of these authors have published five or more works, they also receive additional points for that publication which gives them the second-highest number of points.

b) Special evaluation
Teachers and experts can request that materials falling under categories A.2-A5 and A7.2 be evaluated specially. The Evaluation Committee has the authority to deviate from the evaluation framework (maximum points) as it applies to a given publication. Should a teacher or expert wish for material to be awarded more points than the maximum stated here, supporting arguments must be specifically outlined. In requesting a special evaluation, (s)he must demonstrate that the work in question is exceptionally substantial or appears in an outlet that makes extraordinarily rigorous academic demands. The Evaluation Committee will seek the assistance of specialists in making a special evaluation. The Committee may also take the initiative in seeking for materials to be evaluated specially.

c) Re-evaluation
At five-year intervals, teachers and experts may request that works that have enjoyed extraordinary attention or had an exceptionally great impact within their academic field be re-evaluated. The Evaluation Committee will seek the assistance of specialists in the
academic field in question.

A1-A2

A1 Theses
A1.1 Candidatus or master’s thesis (15 points)
A1.2 Doctoral thesis (30 points)
The name of the university at which the thesis was written must be stated, along with the length and title of the thesis. If an individual has two theses on the same level, both are evaluated for points. If articles, books or other materials based on a thesis are published, these additional publications are evaluated separately in the appropriate category.

A2 Books
A2.1. Peer-reviewed publications by the world’s most respected academic publishers (up to 100 points)
Books published by the world’s most respected academic publishers. The following publishers fall under this category:
Cambridge University Press
Elsevier
Harvard University Press
John Wiley & Sons (incl. Blackwell Publishing)
Kluwer/Springer
Oxford University Press
Taylor and Francis (incl. Routledge)
The above list is not comprehensive. An author whose work has been published by an academic publishing company of the same calibre (including, for example, the foremost publishers within a specific and narrow academic field) may request for this item to be evaluated in this category.
A2.2. International peer-reviewed publications and national peer-reviewed publications with an international dimension (up to 75 points)
Peer-reviewed publications from Iceland are considered to have an international dimension if they meet the following requirements:
1) They must satisfy requirements regarding knowledge creation in the international scientific debate of our time.
2) The publisher guarantees access to the work abroad (through commercial agents, for example), so that it is possible to get them on international book purchasing lists.
3) In addition to peer-review, the publishing company attends to scholarly editing.
4) If material is in Icelandic, it shall be accompanied by a summary in English (or another widely spoken language).
A2.3. Peer-reviewed publications (Icelandic or foreign) primarily aimed at the domestic academic community (up to 50 points)
Materials falling under this category include peer-reviewed works published by recognised publishers but aimed primarily at the domestic academic community that do not satisfy all requirements stipulated for category two. Publications in this category must
be based on independent research, but their worth may at the same time consist in their role in dissemination, i.e., they channel international scientific debate and theories into the domestic academic community.

A2.4. Other books (up to 25 points)
Peer-reviewed scientific works primarily aimed at the domestic academic community.

A2.5. Republications (up to 10 points) Republications must involve substantial change or addition to the previous edition to be evaluated for points.

A3. Book chapters
Book chapters fall into four evaluation categories, analogous to the book categories (A2.1-A2.4) listed here above.

A3.1. Peer-reviewed publications by the world’s most respected academic publishers (up to 20 points)
A3.2. International peer-reviewed publications and national peer-reviewed publications with an international dimension (15 points)
A3.3. Peer-reviewed publications (Icelandic or foreign) primarily aimed at the domestic academic community (10 points)
A3.4. Chapters in other books (0-5 points)

A4 Academic articles
A4.1 Article published in ISI-journal with high impact factor or in an A category journal on the ERIH reference index (20 points).
ISI-journals are those international scientific journals documented in the Institute for Scientific Information (ISI) databases under the auspices of Thomson Reuters. Articles in ISI-journals with an impact factor high enough to place them in the top 10% in their category receive 20 points. The European Research Index for Humanities (ERIH) is based on a system of peer assessment and is under the auspices of the European Science Foundation.
A4.2. Other articles in ISI journals, B category articles (ERIH) or articles in journals that receive a grade I rating in journal survey (15 points).
A4.3. C category articles (ERIH) and articles in journals that receive a grade II rating in journal survey (10 points).
A4.4 Articles published in journals that receive a grade III rating in journal survey (5 points).
The supplement (below) includes a list of the criteria that the survey is based on and details as to how a journal is rated on the basis of the results.

A5 Articles in conference proceedings
An international scientific conference must meet the following criteria:
1. Public programme.
2. Participants chosen on the basis of their expert knowledge.
3. A minimum of 15 speakers.
4. A minimum of 5 speakers work abroad.
Summaries and lengthened summaries are not evaluated for points.
A5.1 Article published in distinguished referenced conference proceedings (10 points)
Peer-reviewed articles in conference proceedings in a recognised publication form within the academic field in question.
Such conference proceedings must be accessible in international databases. Examples of such databases include:
IEEE Xplore: http://ieeexplore.ieee.org
MSME digital store: http://store.asme.org/
SPIE: http://spie.org
ASCE: http://www.asce.org
ACM Digital Library: http://portal.acm.org
This list is not comprehensive. Evaluation Committees evaluate individual cases as they come up.

A5.2 Article in other conference proceedings (3-5)
Conference proceedings are to be peer-reviewed. When awarding points, procedures for peer-review are taken into account.

A6 Lectures and posters
A6.1 Plenary lecture or keynote address at international scientific conference (5 points).
A6.2 Public lecture by invitation at university abroad (3 points).
A6.3 Lecture at international conference (3 points).
A6.4 Lecture at domestic conference (2 points).
A6.5 Lectures at academic symposiums, seminars or meetings for professional groups (1 point).
A6.6 Poster at international conference (2 points).
A6.7 Poster at domestic conference (1 point).
The programme for a conference or meeting must be on hand. A letter of invitation must accompany the performance report in the case of a plenary lecture or keynote address (A6.1) or a public lecture by invitation at a university abroad (A6.2). A plenary lecture at a very large conference (more than 1,000 participants) may be evaluated for 10 points.
Points for lectures and posters are given to the individuals presenting them.
Teachers and experts do, however, receive points if a lecture/poster is held/introduced by their student. Points are calculated according to rules for division of points (see above) as if there were two authors.
When a teacher or expert makes an unusually high number of contributions to the same conference or during a single year, the number of points may be capped should it be deemed that there are grounds to do so.

A7 Editorial work on academic publications
Scholarly editing of academic publications implies editorial work that builds on an editor’s expert knowledge. Copy-editing thus does not fall under this category.
A7.1. Editor of an academic journal (3-6 points/issue)
Only scholarly editing of peer-reviewed journals is evaluated for points. Three points are awarded for editorial work that chiefly involves making decisions regarding publication of material. If scholarly editing also relates to the content and treatment of individual
articles, criticism and scholarly commentary, three additional points are awarded. Should the same individual be responsible for both tasks, (s)he thus receives 6 points per issue. For journals in categories A4.1 and A4.2, an individual receives points for a maximum of three issues per year. For journals in category A4.3, an individual receives points for a maximum of two issues per year and for one issue per year for those in category A4.4.

A7.2. Book editor (5-20 points)
Editing of peer-reviewed books only. The assumption is made that for such books, the editor both makes decisions regarding material published and attends to the editing of individual book chapters. Evaluation also takes into account the scope of the book and its categorisation as outlined in A2. A request may be made for special evaluation for very extensive editorial work, for example when a large anthology with material by a number of scholars is at issue, for example in books re-examining wide-ranging fields within the sciences where the editor takes an active role in this mapping process, i.e., through the organisation of the publication, written introductions to individual book sections, and so on.

A8 Reports, reviews and translations
A8.1 Reports (0-3 points)
Evaluation is based on the scope of data solicitation, originality and contribution to new knowledge. Reports include published academic reports, working papers, university publication series with publication numbers and reports falling under legal deposit legislation (for example having an ISBN number). Reports can also be evaluated for service points (see D4). Reports that are the product of service research generally fall under category D. Evaluation of reports takes in general into account whether the material could, should there be reason to do so, be published in a platform that makes rigorous academic demands.

A8.2 Reviews (1-3 points)
Points are given for reviews in peer-reviewed journals where all published material undergoes a scholarly editing process. Reviews involve scholarly discussion and criticism: they are not short write-ups on the material content of individual books or other publications.

A8.3 Translations of academic articles, book chapters and other short writings (0-5 points)
A8.4 Translations of academic books (10-25 points)
Translations of material of high academic value that fall under the field of study of the individual in question. Republications of works in other languages are valued at up to 10 points. These are translations of academic writings and scientific work published at an international level. Should such a publication bring with it significant academic acclaim, it may be awarded more points in accordance with the provision on special evaluation.

A9 Curriculum design for preschools, primary schools and secondary schools
A9.1 Curriculum design (0-10 points) Teaching materials for preschools, primary schools and secondary schools may be evaluated under this category should they meet with the
following requirements:
1. The subject matter of these teaching materials falls within the author’s area of research.
2. Teaching materials show clear signs of primary research having taken place.
3. Teaching materials are peer-reviewed by experts and published by respected publishers.
4. The goals, methods and process of curriculum design are stated in the associated materials, for example in teaching directions or in an accompanying statement by the author.
5. The teaching materials represent a considerable body of work.
Should these requirements not be met, teaching materials are evaluated for service points.

**A10 Innovation and knowledge transfer**
To receive points for innovation and knowledge transfer, there must be demonstrable connections to research, originality and knowledge innovation. Art creation in an academic context also falls under this category. Innovation and knowledge transfer may also be evaluated for service points.

*A10.1 Start-up company, design, innovation and knowledge transfer (0-20 points).*
Evaluation is based on conventional evaluation criteria, cf. the introduction to these rules. Knowledge transfer is evaluated for points when the establishment of a company or negotiation of a contract involves the release of new knowledge or scientific innovation. The effects that company operations have had on university activities and its relationship with the university are also taken into account. In general, start-up companies are thus not evaluated during the first 5 years of operation. An explanatory report shall accompany the performance report.

*A10.2 Software (0-20 points).*
Software is only evaluated if it involves knowledge creation and/or new software design solutions. The release format is either software distributed as merchandise or open source software.

*A10.3 Psychological tests (0-5 points).*
To be evaluated, tests must be published and accessible and involve research work not appearing in any other form.

*A10.4 Legislative bills (2 points as a general rule).*
Only those bills that involve primary research and research work that has not appeared in any other form are evaluated. Work on a bill not meeting the above requirements may be evaluated for service points. Drafting a bill may in some instances involve considerable research work. In such cases, up to 10 points may be awarded for a bill. An author must specifically make a request for special evaluation, stating his/her reasoning. Evaluation of bills takes in general into account whether the material could, should there be reason to do so, be published in a platform that makes rigorous academic demands. Where the author(s) of a bill are not identified in the bill itself, confirmation of authorship must accompany the performance report.

*A10.5 Patents (10-15 points)*
Only published patents are evaluated for points. Ten points are given for a published patent application and an additional 5 if the patent is granted. No points are given for
republished patents or patent applications.

A10.6 Development work in schools and other institutions (0-10 points)
Development work in schools that builds on research by the scholar in question falls under this category. Development projects in this category involve the transfer of new knowledge within the domain of school operations and are evaluated upon their completion. A project description and implementation plan must be at hand. The connections between the development project and research by the scholar in question must be outlined. When evaluating a development project, the above factors are taken into account as well as how extensive the project is and whether evaluation by peers has taken place. Should a development project not meet the requirements outlined above, it may be evaluated for service points.

Analogous knowledge transfer within other institutions may also be evaluated in this category, provided it satisfies requirements commensurable to those outlined above.

A10.7 Innovation in the arts (0-40 points)
To receive points for art creation, teachers must attend to regular instruction relating specifically to art creation and works of art (including literature). Teachers are not limited to individual art forms in this respect (a visual arts teacher can, for example, receive points for a theatre play or concert performance).

A10.7.1 Large solo exhibition at a recognised art gallery endorsed by an arts council (10-30 points)
Large solo exhibitions of visual and design pieces held at recognised art galleries endorsed by an arts council fall under this category. Such galleries include the National Art Gallery of Iceland, Reykjavík Art Museum, Kópavogur Art Museum, ASÍ Art Museum, Hafnarborg – the Hafnarfjörður Centre of Culture and Fine Art, Living Art Museum, Nordic House of Reykjavík and Akureyri Art Museum.

A10.7.2 Solo exhibition or participation in a group exhibition at a recognised art gallery or at an international arts and culture festival (0-15 points)
Solo exhibitions of pieces not previously shown fall under this category. Recognised art galleries are the same as described above.

A10.7.3 Extensive original musical composition or theatre play publicly performed by recognised artists (10-30 points)
Musical compositions falling under this category include operas and orchestra pieces performed in a public concert hall (the Icelandic Opera, the National Theatre of Iceland, Salurinn (Kópavogur Concert Hall) or the University Cinema). The composition should be of a length to fill an entire programme.

A10.7.4 Original musical composition or theatre play publicly performed by recognised artists (0-10 points)
Shorter compositions performed by recognised artist fall under this category. Points are not awarded more than once for a given composition.

A10.7.5 Musical performance/dramatisation at international arts and culture festivals, at public subscription concerts or at a recognised theatre (0-15 points)
International arts and culture festivals include the International Electronic and Computer Music Festival, Nordic Music Days and other festivals abroad. Public subscription concerts include the Icelandic Symphony Orchestra concert series and Tíbrá concert
series in Salurinn, the Kópavogur Concert Hall.

A10.7.6 Concert or lengthy dramatisation (0-10 points)
Solo performances or artistic direction of orchestra, choir or chamber music ensemble for concert under the auspices of a respected concert holder or cultural institution. Concerts held in connection with or under the auspices of international arts and culture festivals, respected concert holders or respected cultural institutions generally receive more points than concerts held independently by the individual in question. Concert premieres receive more points as a general rule. Instrumentalists may also be awarded points for participation in a chamber music ensemble or important role in a musical performance held by recognised parties, such as the Icelandic Symphony Orchestra, even where a solo performance is not at issue, as they have been selected to perform on the basis of their artistic ability. Such happenings are not, in general, evaluated for points unless they amount to more than two annually. Respected concert holders and cultural institutions include: Icelandic Symphony Orchestra; Salurinn, the Kópavogur Concert Hall; Icelandic Opera; Skálholt Summer Concerts Festival; Reykjavík Arts Festival; Dark Music Days; Reykjavík Chamber Orchestra; Caput; Musica Antiqua; Gerðuberg; University Concerts; Association of Icelandic Musicians; and Music for All.

A10.7.7 Released recordings (0-10 points)
Recordings broadcast by recognised media or released by recognised record labels. Recording and broadcasting at the initiative of a respected medium, i.e., the Icelandic National Broadcasting Service. Recognised record labels include: Association of Icelandic Musicians, Ísdiskar, Iceland Music Information Centre, Jazzís, Klassís, Bad Taste, Skífan, Stöðin Inc., Thule Musik, Tónafloð or known record labels abroad.

A10.7.8 Literary texts. (0-10 points)
Poems, short stories, short plays and other short literary texts.

A10.7.9 Books. (10-40 points)
Works of fiction, poetry anthologies, short story collections, lengthy plays and other literary works of artistic value. Previous publication of individual sections of a book is taken into consideration, as is whether the staging of the play has already been evaluated for points.

A10.7.10 Translations of book chapters and other short texts of artistic value (0-5)

A10.7.11 Translations of books of artistic value (10-25)

A10.7.12 Design, curating or directing work (0-10 points)
Organisation and management of exhibition or show for art gallery or theatre. Design work here implies artistic design of work created for specific needs and contexts, where a number of factors come into play, such as environment, utility and beauty. Individual visual and design pieces are not evaluated for points unless they are designed for specific contexts, where the expertise of the individual in question in designing the piece receives the acknowledgement of other parties such as an arts council or selection committee.

A11 Citations
Citations in the ISI databases (Science Citation Index, Social Science Citation Index and Arts and Humanities Citation Index) are evaluated as follows:
First 10 citations: 1 point/year
Next 20 citations 0.5 point/year
Citations exceeding 30: 0.1 point/year
Citations exceeding 2,000: 0.05 point/year
A request may be made for citations in peer-reviewed books and journals not listed in the ISI-databases to be evaluated. For such publications, the academic in question must provide confirmation that his or her work has been cited.
Points for citations are not paid out of productivity evaluation funds, such as the Writing and Research Fund.

A12 Grants from competitive funds (0-20 points/year)
Total value of grants from parties outside the university in question. Only those grants entered in the accounting system of the university in question or affiliated institutions are evaluated. The project manager or coordinator for a grant application also receives points unless agreed otherwise with grant recipient. Should the grant come from an international competitive fund, points double from what is listed below, but the maximum (20 points) remains unchanged.
Points:
1 point for 0.5-1.999 million ISK/year
2 points for 2-3.999 million ISK/year
3 points for 4-6.999 million ISK/year
4 points for 7-9.999 million ISK/year
One point is given for every additional 10 million ISK/year to a maximum of 20 points/year.
Points for grants are not paid out of productivity evaluation funds, such as the Writing and Research Fund.

Supplement on the categorisation of Icelandic journals
When evaluating an Icelandic journal (cf. A4.2 to A4.4), emphasis is placed on procedure at the journal being in keeping with accepted practices at an international level. To this aim, categorisation uses as a frame of reference those requirements used by Thomson Reuters in the database indexing of international scientific journals. These requirements cover factors such as peer review, whether articles are accompanied by English summary, publication frequency, distribution and submission rejection rate. These 19 requirements are divided into imperative and desirable requirements.

Three points categories of Icelandic journals:
- Articles appearing in superior journals that meet the first 17 requirements listed below receive 15 points.
- Articles appearing in journals that meet the first 15 requirements receive 10 points.
- Articles appearing in journals that meet a minimum of the four imperative requirements listed at bottom receive 5 points.
**Imperative requirements:**
1. Publication of previously unpublished results.
2. Editor and editorial board have pursued higher education in the academic field.
3. Editor or scholarly editorial board review submissions and reject or accept them.
4. Peer review. Articles submitted are never published without anonymous peer review by two or more experts in the area in question.
5. Peer review relates to material content.
6. Peer-reviewed articles specifically marked as such where non-peer-reviewed material is also published.
7. Rejection rate of at least 15% of submissions (calculations based as a rule on the previous three years).
8. Regular, planned publication frequency (5 years taken into account as a rule).
9. Icelandic summary.
12. Descriptive article titles.
13. Satisfactory bibliographical information for all citations.
14. Author addresses satisfactorily well indicated.
15. A minimum of 10% of authors with peer-reviewed articles from outside the university in question.

**Desirable requirements:**
17. Subscribers outside of Iceland or web access.
18. Journal indexed in international databases.

**Imperative requirements:**
1. Publication of previously unpublished results.
2. Editor and editorial board have pursued higher education in the academic field.
5. Comments from editor or editorial board relate to material content.
6. Peer-reviewed articles specifically marked as such where non-peer-reviewed material is also published.

**B. Teaching**

**B1. Teaching experience**

*B1.1. Tenured instructor (senior lecturer, assistant professor, associate professor or professor), full time* 10 points/year

*B1.2. Non-tenured lecturer* 2 points/year for each course taught, to a maximum of 6 points/year

*B1.3. Teaching techniques course* 0-2 points

Teaching in academic mobility programs is evaluated as sessional teaching (B1.2).
B2. Teaching materials
B2.1. Short booklets or teaching materials on the Internet 0-3 points
B2.2. Extensive teaching materials, textbooks 5-60 points
Preliminary versions of teaching materials may be evaluated under B2.1 but fall under B2.2 when published in final form.

B3. Supervision of students and thesis opposition
B3.1. Master’s thesis 2-4 points
B3.2. Doctoral thesis 10 points
B3.3. Member of advisory committee 3 points
B3.4. Thesis opponent 3 points
Points are granted for supervision of graduate students upon completion of the thesis. Students’ names and their thesis titles must be stated. The number of points for a master’s thesis depends on its size. Individuals receive 2 points for supervision of a thesis for fewer than 50 ECTS credits, 3 points for a thesis for 50-70 ECTS credits and 4 points for a thesis for more than 70 ECTS credits.

4. Innovation in teaching 2-10 points
Innovation in teaching can be evaluated for points, such as organising and defining a new study programme, organising new courses or reorganising existing ones, defining new courses, developing teaching methods or designing a project database. A report detailing the nature of the work must accompany an application for evaluation of teaching innovation.

C. Administration
Points are awarded for administrative positions within the university as outlined below. The scope of the position (turnover, number of employees, students, etc.) determines the number of points awarded for C1, C4, and C5.
C1. Department chair 5-10 points/year
C2. Chair of University Council works committee and standing evaluation committee of Academic School 10 points/year
C3. Chair of principal works committees of Academic Schools 5 points/year
C4. Director of research institute 5-15 points/year
C5. Dean/Head of faculty 25-50 points/year
C6. Dean of academic school 75 points/year
C7. Rector 100 points/year
C8. Assistant Rector 50 points/year
C9. Member of University Council 5 points/year
C10. Member of committee under the auspices of the University Council or the Rector 2 points/year
C11. Chair of committee under the auspices of the University Council or the Rector 3 points/year
C12. Member of evaluation committee for position at university 2 points
Academic positions are evaluated for points only after the term is completed. Two points
are awarded for every evaluation committee as per C12 to a maximum of 10 points/year.

D. Service

D1. Organisation of international scientific conference (2-10 points)
Chair/membership in conference committee.

D2. Evaluation work within public sector (0-2 points)
Membership in public evaluation committees, demarcated fixed-term projects.

D3. Committee or board membership (0-2 points)
Outside university in question.

D4. Advisory opinions and reports (0-5 points)
Reports published without formal peer review or unpublished reports and advisory opinions that are released for or under the auspices of parties outside the university in question. The Evaluation Committee must have access to a report for it to be evaluated for points.

D5. Software (0-10 points)
Software must involve applied software development and be distributed nationally or abroad either as merchandise or open source software.

D6. Educational materials for the public (0-20 points)
These must represent the composition and/or dissemination of material that the individual in question undertakes because of his/her academic expertise.
1. Books (general) 0-20 points
2. Translations (general) 0-10 points
3. Editing of non-scholarly books and journals 0-6 points
4. Article in non-scholarly journal 0-5 points
5. Review or critique in the media 1 point
   (maximum: 10 points/year)
6. Newspaper article (0-3 points) or composition of item appearing in other media 0-10 points
   (maximum: 10 points/year)
7. Speech at seminar or symposium aimed at the general public 1 point
8. Dissemination of information and advisory activity 0-10 points
   (maximum: 10 points/year)

D7. Start-up company (0-50 points)
Start-up companies and license agreements with parties outside the university in question. After operations have been going on for some time (e.g., after 5-10 years), the company or agreement may be re-evaluated for up to 50 additional service points. In making the evaluation, employee and student participation, ownership and the visibility of the university in question in connection with the project are taken into consideration.

D8. Grants from competitive funds (0-20 points)
Total value of grants from parties outside the university in question. Only those grants that go through the accounting system of the university in question or affiliated institutions are evaluated. The project manager or coordinator for a grant application also receives one point unless agreed otherwise.

Points:
1 point for 0.5-1.999 million ISK/year
2 points for 2-3.999 million ISK/year
3 points for 4-6.999 million ISK/year
4 points for 7-9.999 million ISK/year
One point is given for every additional 10 million ISK/year to a maximum of 20 points/year.

E. Former employment (for initial evaluation)
Former employment in positions that fall outside these rules is evaluated if it is in the teacher’s or specialist’s field of expertise and relevant to the current appointment.
Work experience: 10 points/year
Work experience points are taken into consideration when making initial wage bracket placements for new appointments. Points for former employment in addition to points for teaching are not to exceed 10 points annually.

F. General
Specialists, academics and scientists with a 40% research requirement shall be awarded 10 points for work experience per year in accordance with the 10 points per year awarded to instructors with a 40% research requirement for teaching experience (see B1).
Specialists, academics and scientists with a 60% research requirement shall receive 7 points per year, while specialists, academics and scientists with an 80% research requirement receive 3 points per year. These points in addition to points awarded for overtime instruction shall not, however, exceed a total of 10 points per year.
The Evaluation Committee has the authority to evaluate projects that do not fall within the scope of these rules for points should a well-founded request or recommendation be put forward, in particular for work inside or outside the university benefiting the scientific community in a broad sense and extensive promotion or education work aimed at the general public.
Exceptions may be made to these rules should there be special cause to do so.