Dissertation

# Talent Proof

## Selection Processes in Research Funding and Careers

Pleun van Arensbergen

Rathenau Instituut

The **Rathenau Institute** promotes the formation of political and public opinion on science and technology. To this end, the Institute studies the organization and development of science systems, publishes about social impact of new technologies, and organizes debates on issues and dilemmas in science and technology.

**Talent Proof**

Selection Processes in Research Funding and Careers

Rathenau Instituut
Anna van Saksenlaan 51

Postal address:
Postbus 95366
2509 CJ Den Haag

Telephone: +31 (0)70 342 15 42
E-mail: info@rathenau.nl
Website: www.rathenau.nl

Publisher: Rathenau Instituut
Lay-out: Boven de Bank, Amsterdam
Coverphoto: Shutterstock
Printing: Drukkerij Quantes, Rijswijk

This book is printed on FSC certified paper.

ISBN/EAN: 978-90-77364-57-4

Preferred citation:
van Arensbergen, P., Talent Proof: Selection Processes in Research Funding and Careers,
Den Haag, Rathenau Instituut 2014

This dissertation is based on research the author has done as an employee of the Rathenau Instituut. This text is the author's PhD thesis. The views in this dissertation are not necessarily the same as those of the Rathenau Instituut.

VRIJE UNIVERSITEIT

# Talent Proof

## Selection Processes in Research Funding and Careers

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Sociale Wetenschappen
op maandag 22 september 2014 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Pleuntje van Arensbergen

geboren te Venray

promotor: Prof. dr. P.A.A. van den Besselaar
co-promotor: Dr. I.C.M. van der Weijden

Thesis committee:
Prof. dr. B. Beersma (Vrije Universiteit)
Prof. dr. S.E. Buitendijk (Leids Universitair Medisch Centrum)
Dr. B.J.R. van der Meulen (Rathenau Instituut)
Dr. C.J. Vinkenburg (Vrije Universiteit)
Prof. dr. P.F. Wouters (Leiden Universiteit)

Voor oma, zo dichtbij

**Little Dandelion**

Brave little Dandelion
Fast falls the snow,
Bending the daffodil's
Haughty head low.
Under that fleecy tent
Careless of cold,
Blithe little Dandelion
Counteth her gold.

*Helen Barron Bostwick*

# Content


**1 Introduction** **13**
1.1 Current debates in academia 13
1.2 General scope of this dissertation 14
1.3 Changing relations within academia 14
1.4 Academic career structures in the Netherlands 16
1.5 Focus and outline of this dissertation 18
1.6 References 20

**2 Different views on scholarly talent:**
**What are the talents we are looking for in science?** **19**
2.1 Introduction 19
2.2 Research questions 20
2.3 Theoretical background 20
2.4 Data and methods 23
2.5 Results 24
2.6 Conclusions and implications 34
2.7 References 36

**3 The selection of talent as a group process**
**A literature review on the social dynamics of decision-making in grant panels** **38**
3.1 Introduction 38
3.2 Focus of this review 39
3.3 Methodology 40
3.4 Panel review of grant applications 41
3.5 Peer review as social interaction 44
3.6 The organizational context of panel review 50
3.7 Conclusions and implications 52
3.8 Future Research 53
3.9 References 54

**4 Academic talent selection in grant review panels** **59**
4.1 Introduction 59
4.2 Theoretical background 60
4.3 Data, research questions and methods 62
4.4 Results 66
4.5 Conclusions and discussion 81
4.6 References 82

**5 Gender differences in scientific productivity: a persisting phenomenon?** **86**
5.1 Introduction 86
5.2 Research question 89
5.3 Materials and methods 89

5.4 Gender differences 91
5.5 Changing gender differences? 93
5.6 A more detailed view on specific disciplines: psychology and economics 95
5.7 Conclusions and discussion 97
5.8 References 97

**6 Determinants of success in academic careers 101**
6.1 Introduction 101
6.2 Research questions 101
6.3 Data and methods 102
6.4 The case 104
6.5 Findings 106
6.6 Conclusions and discussion 116
6.7 References 117

**7 Conclusions and general discussion 121**
7.1 Summary of most important results 121
7.2 Further reflection upon the current selection system and its accompanying criticism 125
7.3 Implications for practice to uphold and stimulate talent 128
7.4 Implications for future research 130
7.5 References 131

Nederlandse samenvatting 132
Dankwoord 138

# 1 Introduction

For decades there has been an ongoing discussion about quality in science. Questions related to quality have been the subject of a long research tradition, i.e. what is scientific quality and how to evaluate and improve it (e.g. Butler, 2007; Gibbons and Georghiou, 1987; Hemlin and Montgomery, 1990; Wolff, 1970)? The same questions have been asked with regard to the societal quality of research (e.g. De Jong, Van Arensbergen, Daemen et al., 2011; Donovan, 2007; Spaapen, Dijstelbloem and Wamelink, 2007), and to human capital and academic talent (e.g. Van den Brink, Fruytier and Thunnissen, 2012). Due to both the complexity and relevance of these concepts the discussion is still ongoing.

## 1.1 Current debates in academia

Recently there has been rising turmoil within the academic community. While not everyone agrees, many academics worldwide are expressing more and more criticism and concerns about the functioning of contemporary academic systems. The current academic incentive system is criticized as overemphasizing the need to publish in high-profile journals. This is said to encourage rapid submissions (Collins and Tabak, 2014), and to prevent data sharing and replication of previous studies (*The Economist,* October 19th 2013a). The principle of self-correction by means of verification is generally considered to be an important cornerstone of science, but is considered under threat by the current academic climate. As research aiming to verify earlier studies is generally considered not to be highly pioneering or interesting, it is often difficult to get funded and published. Replication studies are claimed to do little to advance a researcher's career (*The Economist* October 19th 2013b).

Universities are criticized as supposedly transforming themselves into 'publication factories', where publishing is strongly overrated and teaching underrated (Dijstelbloem, Huisman, Miedema et al., 2013). In the current debates, mentoring and teaching are emphasized to be very important, as they involve educating new generations of researchers, but to be lacking in the assessment of researchers for hiring, funding or job promotion purposes (Fang and Casadevall, 2011). Whereas mentoring and teaching achievements are perceived as hardly contributing to career advancement, publishing in prestigious journals is claimed to do so, even to such an extent that publishing is criticized as having become the ultimate goal of academic labour these days.

Furthermore, it is claimed there is too much focus on scientific excellence and star scientists, and too few career opportunities for young researchers. Growing numbers of PhD students and postdoctoral researchers are trained at universities. They are alleged to do significant work without much perspective on an academic career (Halffman and Radder, 2013). Early career grants are available, but only for what is seen as the small group of excellent young researchers. This relates to the severe competition characterizing modern science. Academia is even described as a 'hypercompetitive' environment: "While some competition is inarguably good for science, excessive competition is demoralizing, destructive, and counterproductive" (Fang and Casadevall, 2011, p. 897).

Lately, several alignments have been established which revolt against the current situation and plead for change. For example, the San Francisco Declaration on Research Assessment (DORA)

was initiated in 2012 by the American Society for Cell Biology together with a group of editors and publishers from scholarly journals. This group predominantly aims at reducing the dominant role of journal impact factors in research assessments. These factors are said to have become a powerful proxy for scientific value and to be widely misused in the assessment of individual academics. According to DORA, the focus of assessments should be on the content of primary research papers instead of where they are published (am.ascb.org/dora). In the Netherlands similar alignments include the Concerned VU group (*Verontruste VU'ers*, 2012) and Science in Transition (Dijstelbloem, et al., 2013). The Concerned VU group, consisting of academic and administrative staff of the VU University Amsterdam, claims that the university is too strongly steered by economical motives and too strongly managed by managers who lack affinity with scientific work or ethics. The movement called Science in Transition pleads for more appreciation of the added value of science for society: decisions regarding knowledge production should not only be in the hands of scientists, but should involve societal stakeholders as well.

## 1.2 General scope of this dissertation

Not everyone agrees on the claims described above and although some of them are broadly supported worldwide, in general they lack a sound empirical basis for thorough analyses. This dissertation is about academic careers, and more specifically, talent selection regarding early career researchers in the competitive context of grant allocation within the Netherlands. It provides empirical material which may offer the opportunity to test several of the claims related to the overpowering role of publications, the negative consequences of competition and the lack of career opportunities for early career academics. Through analyses of current evaluation and selection processes, we will create a better understanding of which criteria are used and the weight assigned to publications as criteria for talent assessment within the young generation of academics. Furthermore, we look at the potential consequences of competition and at career opportunities for early career researchers.

We based our research on data from the Netherlands as this is an interesting case for several reasons. First, the Netherlands is one of the better performing Higher Education systems. It has a high position in various international university ranking tables and it has high average publication and citation scores. Second, the personal funding program de *Vernieuwingsimpuls* of the Dutch Research Council (NWO) is considered to be highly successful. Internationally it is often seen as a model example, proven by several similar funding programs being set up in the recent past. And third, in the Netherlands the role of talent management is growing in Dutch academia. In the past decade, many talent programs and initiatives were implemented to stimulate, attract and develop academic talent. For a better understanding of the current academic career and selection system in the Netherlands, including the focus on talent and excellence, the most important changes within higher education in the past few decades leading to the current system, are briefly described below.

## 1.3 Changing relations within academia

Since the 1980s, overall R&D expenditures in most OECD countries more than doubled, and within the higher education sector, even tripled. Growing investments in higher education have resulted in substantial growth in student numbers, causing a transition from elite to mass education (Vincent- Lancrin, 2006). There was also a strong increase at the PhD level: OECD countries

showed a 40% increase in doctoral graduates between 1998 and 2006 (Auriol, 2010). The 'massification' of higher education has subsequently affected the academic profession, leading to a rise in the number of academic staff on all levels. There has been a strong worldwide increase in the category of postdoctoral researcher in particular. For example, in the United States, the number of postdocs in 2010 increased by about 30% compared to 2005 and by 250% when compared to 1979 (Einaudi, Heuer, and Green, 2013). In the Netherlands within this similar more recent period (2005-2011), the number of postdocs increased by around 40% percent (De Goede, Belder and De Jonge, 2013), compared to rises of 11% for assistant professors, 7 % for associate professors and 17% for full professors (VSNU/WOPI).

At the same time due to increasing pressure towards more efficient use of public resources, 'New Public Management' was introduced. The government's role in higher education diminished, empowering European universities as institutional actors (Goedegebuure and Meek, 1997; Leisyte, 2007; Vincent-Lancrin, 2006). This responsibility incited universities to develop more management and evaluation tools. Assessment has become an ongoing activity. The object of assessment procedures varies from individual researchers to whole institutes, from teaching to research performance, and from ex-post to ex-ante. Academics currently face many more moments of evaluation than in the past, in line with the emerging audit culture as also exists in broader society (Strathern, 2000). Nowadays academics are continuously monitored in terms of their performance not just in situations related to recruitment or promotion. Besides increased emphasis on output control and performance measures, decisions with regard to budgets, staff and strategies, previously mainly managed by public authorities, have to a large extent been transferred to universities. Nowadays most European universities, although there are differences between countries, have the authority to manage their budgets, implement their scientific and organizational strategies and policies, and over recruitment and staff management. They determine the size of their staff, the type of positions (e.g. academic, administrative, junior, senior, temporary, tenure) and the timing and procedure of recruitment. This has led to the situation in which universities have more control over academic careers (Musselin, 2013a). However, these transitions have not only changed the balance of power in university management, but also for academic elites (Musselin, 2013b). Full professors and academic staff play a crucial role in internal university policies. They are key actors in the execution and management of academic organizations (Thunnissen and Fruytier, 2014).

At the same time, as rising numbers of students and academic staff has not been matched by an increase of public funding, universities have become more dependent on external funding and more market oriented. Although in most European countries the amount of government funding is generally still increasing and it is a major funding source for academic research, funding has become more project- orientated (Lepori, Van den Besselaar, Dinges et al., 2007) and other sources like private funding gained importance (Vincent-Lancrin, 2006). The shift from block funding towards project funding has contributed to increased competition and a changing role for intermediary organizations. Instead of government funding being allocated to the entire university or research institute, it is to an increasing extent allocated via intermediaries directly to individual researchers or research groups (Lepori, et al., 2007). Universities are encouraged to compete for external funding from industry, national and European research councils, or charity foundations, also as part of a deliberate strategy to strengthen ties with society and stakeholders (Mouwen, 2000).

The increases in academic staff supply and the changes in the funding landscape have required universities to adjust their career system and recruitment policies. They have reduced the share of tenured staff in order to increase their flexibility and adaptive power towards varying external demands and circumstances. This has led to a rise in temporary staff, but also to a wider variety of academic positions, e.g. research project staff and teaching-only staff (Huisman, De Weert and Bartelse, 2002; Santiago and Carvalho, 2008). Until recently, academic careers could be described in terms of two stages: a training and temporary position stage and a permanent position stage. At the start of a career, people were predominantly engaged in training and developing academic skills, followed by a selection. Those who passed were appointed to fixed-term contracts. The second period was characterized by access to permanent positions. Nowadays, this clear structure is no longer generally valid, as people more often work in several temporary positions successively, without getting a permanent contract (Enders and Musselin, 2008).

## 1.4 Academic career structures in the Netherlands

As this study is conducted within the context of the Dutch career system, we will now zoom in on the Netherlands. *Figure 1* illustrates the academic ranking structure in the Netherlands, according to position, age and gender. Professors, associate professors and about two thirds of the assistant professors have permanent contracts; postdocs and PhD students generally have temporary positions (De Goede et al., 2013). About 60% of all academic staff is employed on a temporary basis. This percentage includes PhD students, whose number is rising and who all have temporary contracts.[1] Excluding PhD students, the share of temporary academic staff is 40%, still more than twice as high as in the national labour force, where 19% of the employees has a temporary contract (www.vawo.nl).

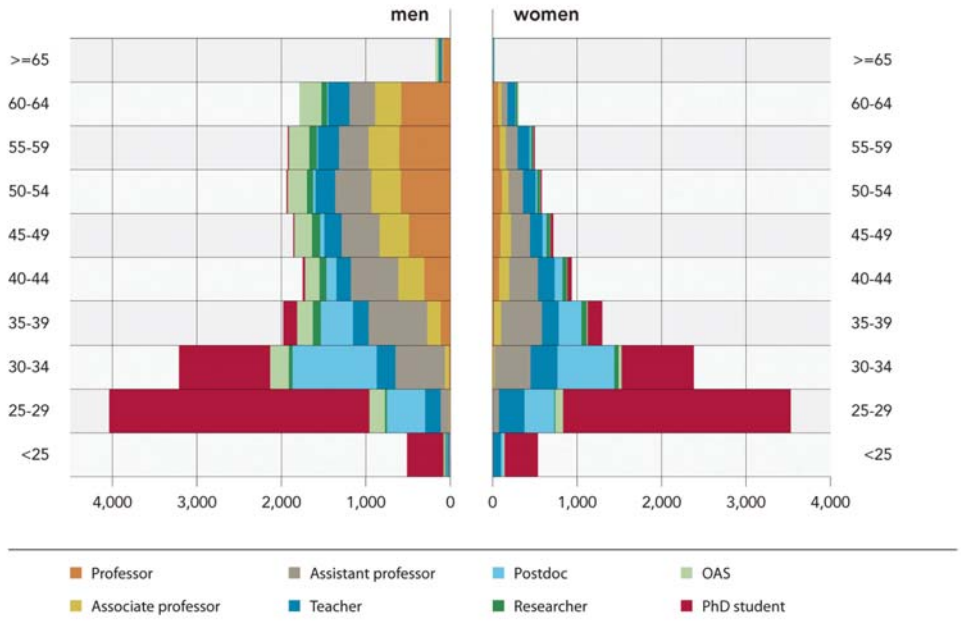
Two main conclusions can be drawn from the figure: strong selection takes place in early career phases and there are gender differences in career structures. The first observation related to *figure 1* is the significance of selection in the early career phases, where the number of positions strongly reduces with each next step upwards.[2] The supply of young academic staff exceeds the academic career opportunities by far. Even though this selection is not unique to the academic sector, it does point out that doctoral education cannot be considered as a mechanism to create new generations exclusively of scholars. Due to the status of PhD candidates' contracts in the Netherlands (formal work contract), PhD positions are generally considered a first step in an academic career. And in line with this general perception, most postdocs aspire to an academic career (Thunnissen and Fruytier, 2014). However, as a growing number of graduates will be employed outside academia, doctoral education to a large extent trains people for the non-academic labour market (Huisman et al., 2002). For example, in the U.S. approximately half of doctoral graduates pursue an academic career (Hoffer, Dugoni, Sanderson et al., 2002), in the

1 In the Netherlands, PhD students are generally employed by the universities, and are included in staff numbers. Besides this type of PhD student, there are external PhD students (e.g. working part time on their PhD while working elsewhere) and scholarship recipients (provided with a stipend or scholarship for their PhD research). The exact size of these two groups is unclear, as they are not registered as such by universities, but is estimated together to be at least as large as the first group.

2 Note that only PhD students employed by universities are included. Although external PhD candidates might differ in their ambitions for an academic career, as they are not included in *figure 1*, the actual selection is stronger than depicted here.

Netherlands this is about 30% (Van Balen and Van den Besselaar, 2007). The large majority will seek a job outside academia.

**Figure 1** Academic ranking structure in the Netherlands according to age, function and sex




Source: VSNU/Wetenschappelijk Onderwijs Personeelsinformatie (WOPI)

**Rathenau Instituut**

Related to the second observation, the figure shows men are still overrepresented in higher positions. The shape of the career figure for women academics looks much more like a pyramid than that of men, where the ranking goes up more evenly. All over the world the share of women is growing for all academic positions, however the general rule still is 'the higher the rank in academia, the lower the number of women' (Brouns, 2000; De Weert, 2001; Timmers, Willemsen and Tijdens, 2010). In the Netherlands the current share of women students and PhD students are 50% and 45% respectively, whereas the share of women professors only slowly moved up to 15%, one of the lowest percentages in Europe (Gerritsen, Verdonk and Visser, 2012). Currently, the largest bottleneck seems to be at the transition from assistant to associate professor (De Goede et al., 2013). The claim that this would be the result of women being less ambitious than men, does not hold as previous studies do not confirm gender differences in ambition (e.g. Dikkers, Van Engen and Vinkenburg, 2010). The stagnation of women in the academic system is considered to be the result of persistent gender disparities in science (Lariviere, Ni, Gingras et al., 2013). More generally, gender disparity can be ascribed to a 'male model' of science, including masculinity of organizational, social and cultural norms within academic organizations (Bleijenbergh, Blonk, Schulte et al., 2008). For example, university standards implicitly prescribe working fulltime and overtime, full availability and flexibility of staff. Furthermore, there is a strong

emphasis on early achievements and mobility, mainly disadvantaging female academics who often have greater responsibilities towards starting and maintaining a family, particularly early on in their career (Bleijenbergh, Benschop and Vennix, 2013; Etzkowitz and Ranga, 2011; Ranga, Gupta and Etzkowitz, 2012).

In order to prevent excellent scholars from leaving academia and due to a growing awareness of the importance of human capital, more and more attention is being paid to talent selection and talent management. More than ever before, universities explicitly aim for attracting and retaining talent. A growing number of European organizations is currently using the Human Resources Strategy for Researchers (HRS4R) to align their policies and practices to the principles of the European Commission's Charter and Code. In acknowledgement of their efforts they can display the HR Excellence in Research logo awarded by the European Commission. With this logo organizations can present themselves as providers of a stimulating and favourable work environment, committed to fair and transparent recruitment and appraisal procedures. Various talent selection and development programs have been implemented so far. For example, following the American example, many European countries currently recruit academics via tenure tracks, offering high potentials a promotion trajectory based on individual performance. The employee is hired on a temporary contract, with the prospect of a permanent contract if (s)he meets certain performance criteria within the first few years. Furthermore there are scholarships for excellent PhD students, mentoring programs for early career researchers, programs promoting equal opportunities for women, and coaching programs stimulating and supporting high potentials in applying for external funding (De Boer and Jongbloed, 2010; Huisman, De Weert and Bartelse, 2002; Neufeld, Huber and Wegner, 2013; Van den Brink, Fruytier and Thunnissen, 2012).

Also with regard to funding worldwide there is an increased focus on talent and excellence. Not only because grants are to a large extent distributed via competition, but also as various countries set up personal funding programs especially directed at talented early career researchers (e.g. the 'Future Research Leaders' grants of the UK's Economic and Social Research Council (ESRC), the Saupere Aude Postdoc grants of the Danish Council for Independent Research (DFF) or the Veni grants of the Netherlands Organization for Scientific Research (NWO)). Although the size of these programs varies, they do take up a significant part of the research councils' budgets (Van Arensbergen, Hessels and Van der Meulen, 2013). The growing importance of these programs, reflected in their budget sizes and application pressures, places these intermediary organizations - peopled with members of the academic community - in a more central position within the academic playing field (Musselin, 2013b).

## 1.5 Focus and outline of this dissertation

In short, in the past decades, the balance in temporary and tenure positions has shifted much more towards temporary positions. Academics have become more dependent on competitive external project funding and individual funding. The relations have altered between the various actors within the academic playing field: the academic staff and their peers, the academic organization including human resource management (HRM) and administration, and the intermediary research councils. All these developments have led to an increase of evaluation practices and occurrences, to intensified competition for funding and academic positions, and finally to more emphasis on talent and excellence, especially in early careers. In this study the main focus is

on the concept of talent and we will examine the process of talent selection within the academic community. The overall aim of this study is to create a better understanding of the process of talent selection. By using different methods and combining various levels of aggregation, we will answer overarching research questions as *what is academic talent* and *how is it selected*?

*Chapter 2* analyses the notion of talent in more detail. Based on interviews with established academics involved in the allocation of personal research grants, we will compare concrete and general views on talent. How do they recognize talent among grant applicants and among academics in their own daily work environment? Furthermore, by analyzing the process of grant allocation we create a better understanding of which criteria are used during the selection process and how final allocation decisions are established within panels.

Since talent selection is increasingly performed by panels, *chapter 3* shows what is known from the literature on this type of panel reviewing and group decision making. This literature review combines peer review studies from the *sociology of science* and *science policy studies* with group dynamics studies from *social psychology* to better understand the dynamics of these selection decisions within panels.

*Chapter 4* is a quantitative study of the talent selection procedure. Through statistical analyses of the reviews of about 900 grant applications, we will show how the followed procedure affects the selection decisions. Furthermore, it shows the importance of the various phases and criteria within the procedure. For example, what is the influence of the external peer reviews and what is the influence of the interview with the candidate? This chapter also answers the question how evident talent is and whether real top talents can be identified. Subsequently it studies the relation between talent selection and gender. Is talent selection gender-neutral and does panel composition matter?

In *chapter 5* we take a closer look into the issue of gender bias. One of the potential factors explaining the underrepresentation of women in higher academic positions is the difference in scientific performance between men and women. Are men outperforming women? By analyzing the publication and citation records of 843 social scientists, we study whether these often demonstrated gender differences still exist in younger generations of researchers, and whether men still outperform women in terms of productivity.

After several chapters on the process of talent selection, *chapter 6* is about the talents themselves. Interviews are conducted with academics who were previously identified as talent, but of whom several have left academia in the meantime. A comparison is made between the careers of talents who stayed in academia and of those who left academia. It is an exploratory study on the importance of various factors deciding for talents to stay within science or to switch to a career outside science.

This dissertation concludes with a general discussion on the current mechanisms of talent selection. After summarizing all chapters and more specifically the main findings regarding the notion and recognition of talent, we discuss the potential impact of the current selection mechanisms on academic talent. What are the potential consequences for the talented academics and for the academic organizations? Finally, the implications for practice and future research are given.

## 1.6 References

Auriol, L., Felix, B. & Schaaper, M. (2010). Mapping careers and mobility of doctorate holders. STI Working paper 2010/1, Paris: OECD.

Bleijenbergh, I., Benschop, Y. & Vennix, J. (2013). Group Model Building to support gender equality change in: Achterbergh, J., Benschop, Y., Hendriks, P. & Van de Ven, A. (Eds.): *Op zoek naar het andere: een liber amicorum voor Hans Doorewaard*. Boom: Meppel: 80-96.

Bleijenbergh, I.L., Blonk, E., Schulte, L., & Van Engen, M. (2008). *Explaining women's careers at a Dutch university: Model building as a method for knowledge elicitation in gender analysis*. 26th International Conference of the System Dynamics Society. System Dynamics Society: Athens, Greece.

Brouns, M. (2000). The gendered nature of assessment procedures in scientific research funding: The Dutch case. *Higher Education in Europe, 25*, 193-201.

Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy, 34*(8), 565-574.

Collins, F.C. & Tabak, L.A. (2014). NIH plans to enhance reproducibility. *Nature, 505*, 612-613.

De Boer, H. & Jongbloed, B. (2010). *Analyse van universitaire jaarverslagen 2009*. Achtergrondstudie voor Rathenau Instituut. Enschede: CHEPS.

De Goede, M., Belder, R. & De Jonge, J. (2013). Facts and figures: Academic careers in the Netherlands. Den Haag: Rathenau Instituut.

De Jong, S., Van Arensbergen, P., Daemen, F., Van der Meulen, B. & Van den Besselaar, P. (2011). Evaluation of research in context - an approach and two cases. *Research Evaluation, 2*0(1), 61-72.

De Weert, E. (2001). Pressures and prospects facing the academic profession in the Netherlands. *Higher Education, 41*(1-2), 77-101.

Dijstelbloem, H., Huisman, F., Miedema, F. & Mijnhardt, W. (2013). Waarom de wetenschap niet werkt zoals het moet, en waar daar aan te doen is. *Position paper*, Universiteit Utrecht.

Dikkers, J., Van Engen, M. & Vinkenburg, C. (2010). Flexible work: ambitious parents' recipe for career success in the Netherlands. *Career Development International, 15*(6), 562-582.

Donovan, C. (2007). The qualitative future of research evaluation. *Science and Public Policy, 34*(8), October, 585-597.

Einaudi, P., Heuer, R. & Green, P. (2013). Counts of postdoctoral appointees in science, engineering, and health rise with reporting improvements. Arlington, VA (NSF 13-334).

Enders, J. & Musselin, C. (2008). Back to the future? The academic profession in the 21st century. In: *OECD Higher Education 2030*. Vol. 1. Paris: OECD.

Etzkowitz, H. & Ranga, M. (2011). Gender dynamics in science and technology: from the "leaky pipeline" to the vanish box. *Brussels Economic Review, 54* (2-3), 131-147.

Fang, F.C. & Casadevall, A. (2011). Reforming science: structural reforms. *Infection and Immunity, 80*(3), 897-901.

Gerritsen, M., Verdonk, T. & Visser, A. (2012). Monitor vrouwelijke hoogleraren 2012. Groningen: Stichting de Beauvoir.

Gibbons, M. & Georghiou, L. (1987). Evaluation of research. A selection of current practices. Paris: Organisation for Economic Co-operation and Development (OECD).

Goedegebuure, L. & Meek, V.L. (1997). On change and diversity: The role of governmental policy and environmental influences. *Higher Education in Europe, 22*(3), 309-319.

Halffman, W. & Radder, H. (2013). Het academisch manifest. Van een bezette naar een publieke universiteit. Krisis: *Journal for Contemporary Philosophy, 3*, 1-21.

Hemlin, S. & Montgomery, H. (1990). Scientists' conceptions of scientific quality. An interview study. *Science Studies, 1*, 73-81.

Hoffer, T. B., Dugoni, B. L., Sanderson, A. R., Sederstrom, S., Welch, V., Guzman-Barron, I., & Brown, S. (2002). *Doctorate recipients from United States universities: Summary report 2001*. Chicago: NORC at the University of Chicago.

Huisman, J., De Weert, E. & Bartelse, J. (2002). Academic careers from a European perspective: the declining desirability of the faculty position. *The Journal of Higher Education, 73*(1), 141-160.

Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C.R. (2013). Global gender disparities in science. *Nature, 504*, 211-213.

Leisyte, L. (2007). University Governance and Academic Research. Enschede: University of Twente.

Lepori, B, Van der Besselaar, P., Dinges, M., Poti`, B., Reale, E., Slipersaeter, S., Theves, J. & Van der Meulen, B. (2007) 'Indicators for comparative analysis of public project funding. Concepts, implementation and evaluation', *Research Evaluation 16*(4): 243-255.

Mouwen, K. (2000). Strategy, structure and culture of the hybrid university: towards the university of the 21st century. *Tertiary Education and Management, 6*, 47-56.

Musselin, C. (2013a). Redefinition of the relationships between academics and their university. *Higher Education, 65*, 25-37.

Musselin, C. (2013b). How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate, *Research Policy*, June 2013.

Neufeld, J., Huber, N. & Wegner, A. (2013). Peer review-based selection decision in individual research funding, applicants' publication strategies and performance: The case of the ERC Starting Grants. *Research Evaluation, 22*(4), 1-11.

Ranga, M., Gupta, N. & Etzkowitz, H. (2012). *Gender effects in research funding: a review of the scientific discussion on the gender-specific aspects of the evaluation of funding proposals and the awarding of funding*. Bonn: Deutsche Forschungsgemeinschaft.

Santiago, R. & Carvalho, T. (2008). Academics in a new work environment: The impact of New Public Management on work conditions. *Higher Education Quarterly, 62*(3), 204-223.

Spaapen, J, Dijstelbloem, H. & Wamelink, F. (2007). *Evaluating Research in Context: a Method for Comprehensive Assessment*. The Hague: COS.

Strathern, M. (2000). *Audit Cultures. Anthropological studies in accountability, ethics and the academy*. London: Routledge.

The Economist (October 19th 2013a). Unreliable research. Trouble at the lab.

The Economist (October 19th 2013b). Problems with scientific research. How science goes wrong.

Timmers, T. M., Willemsen, T. M., & Tijdens, K. G. (2010). Gender diversity policies in universities: A multi-perspective framework of policy measures. *Higher Education, 59*(6), 719-735.

Thunnissen, M. & Fruytier, B. (2014). Het mobiliseren van Human Capital: een overzicht van 25 jaar HRM-beleid op Nederlandse universiteiten. *Tijdschrift voor HRM, 1*, 1-24.

Van Arensbergen, P., Hessels, L., Van der Meulen, B. (2013). *Talent Centraal: ontwikkeling en selectie van wetenschappers in Nederland*. The Hague: Rathenau Instituut.

Van Balen, B. & Van den Besselaar, P. (2007). *Universitaire onderzoeksloopbanen: een verkenning van problemen en oplossingen*. The Hague: Rathenau Instituut.

Van den Brink, M., Fruytier, B. & Thunnissen, M. (2012). Talent management in academia: Performance systems and HRM policies. *Human Resource Management Journal, 23*(2), 180-195.

Vincent-Lancrin, S. (2006). What is changing in academic research? Trends and future scenarios. *European Journal of Education*, *41*(2), 169-202.

Verontruste VU'ers (2012). Dekoloniseer de VU. We zijn geen koekjesfabriek. Retrieved 11 November, 2013, from http://verontrust.wordpress.com/2012/06/11/manifest/.

VSNU, Wetenschappelijk Onderwijs Personeelsinformatie (WOPI). http://www.vsnu.nl/f_c_personeel_downloads.html

Wolff, W.M. (1970). A study of criteria for journal manuscripts. *American Psychologist*, *25*, 636-639.

http://www.vawo.nl/ruim-40-procent-van-de-gepromoveerde-wetenschappers-bij-de-universiteiten-is-in-tijdelijke-dienst

# 2 Different views on scholarly talent: What are the talents we are looking for in science?[1]


## Abstract

In this paper we study the evaluation of talented early career researchers, as done in grant allocation processes. To better understand funding decisions, we studied the grant allocation process in more detail, and compare the notion of talent in grant allocation with more general notions of talent existing in the academic work environment. The comparison is based on interviews with 29 scholars who have experience with identifying talent both in their daily academic work and in the process of grant allocation. Overall there is large agreement on the notion of talent. However, the characteristics ascribed to top talent vary depending on the evaluation context. In grant allocation a narrower talent definition prevails compared to more general evaluation. Furthermore, difficulties arise in the process of panel decision-making, when selection criteria need to be concrete and explicit to enable comparison. Having to choose between many applicants of similar quality makes the selection process liable to subjectivity, arbitrariness and randomness. Despite these uncertainties, grants are ascribed a very high symbolic value. Small quality differences are enlarged into considerable differences in recognition, consequently affecting career opportunities, as they provide academics with both financial and symbolic resources.


## 2.1 Introduction

The quality of higher education and research is strongly connected to the quality of the people working in the academic sector. For excellent science, excellent scientists are needed. Therefore government and universities specifically aim at selecting the best people when investing public resources in education and research. This has led to an increased focus on talent and talent policy, especially within the group of early career researchers. This is a general phenomenon, also clearly visible within the Netherlands, the specific context of this study. In the hiring policies of Dutch universities the notion of talent nowadays has taken a central position. It is becoming a key issue in Higher Education for Human Resource Management (HRM); a management tool that focuses on individual performance (Waring, 2013). Several programs and policy initiatives are currently implemented to attract and stimulate academic talent, e.g. the Tenure Track program, scholarships for excellent PhD students, and mentoring programs for promising female academics (De Boer & Jongbloed, 2010; Van den Brink, Fruytier & Thunnissen, 2012). Furthermore, many programs are directed towards motivating researchers to apply for external grants and to increase the chances of young researchers to acquire external funding. They involve pre-selection processes, encouragement by dedicated mentors, training and supervision of writing grant applications, and improvement of presentation skills (Neufeld, Huber and Wegner, 2013).

The increased focus on acquiring external funding is both a deliberate consequence of changes in funding policy (from institutional towards project funding), as it is a result of the growing pool

1 This chapter has been accepted for publication by Research Evaluation as Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (*forthcoming*). Different views on scholarly talent: What are the talents we are looking for in science?

of competent scholars with academic career ambitions. Government funding used to be predominantly allocated as block funding to the entire university or research institute. Nowadays, in many countries intermediaries like research councils have taken an important position in the distribution of resources, as they allocate an increasing share of government funding directly to individual researchers or research groups (Lepori et al., 2007). Securing external research grants has become a prominent criteria in academic recruitment, evaluation and promotion processes in science (Bloch, Graversen & Pedersen, 2014; De Jonge Akademie, 2010; Laudel & Gläser, 2012; Van Arensbergen, Hessels & Van der Meulen, 2013). Therefore, career opportunities as well as university appointment decisions of individual scientists depend on granting decisions made by external funders, like the European Research Council (ERC) or the Dutch Research Council (NWO). Personal career grants are to an increasing extent considered as a necessary resource in order to further develop an academic career. The real effect of these grants on someone's career is subject of many studies. A recent study (Laudel & Gläser, 2012) on the impact of ERC grants on careers of grantees, showed that several organizations responded to the reputation of ERC grants by promoting (mainly ERC starting) grantees or by offering them tenure. At the same time the grants were found to only play a minor role in promoting inter-organizational mobility. Most of the grantees already worked in the best possible environments, and/or were settled with their families which constrained them to move. A Danish study showed that grant recipients from the Danish council for Independent Research have a higher probability of becoming a full professor (16%) compared to rejected applicants (9%). Also grant recipients stressed the central role of grants in facilitating subsequent collaboration with leading researchers in their field and in establishing their own positions in research communities (Bloch et al., 2014).

A successful personal funding instrument of the Dutch Research Council is the Innovational Research Incentives Scheme (*Vernieuwingsimpuls*). Gerritsen, Plug and Van der Wiel (2013) recently studied the effect of these grants and found that they indeed increase one's chance of a successful academic career. Comparing applicants with about the same priority scores, they found that grantees are more likely than rejected applicants to stay in academia, receive follow-up grants and become a full professor. These results show that the allocation of personal career grants is an important context of talent selection. To study the notion of talent within science, the process of grant allocation is therefore included and investigated in more detail.

## 2.2 Research questions

In this study we investigate the process of talent selection in more detail with a twofold aim. On the one hand we want to create a better understanding of the notion of academic talent. What are the talents the academic community is looking for, both within academia in general as within grant allocation procedures? On the other hand we want to open up the black box of grant allocation by scientific panels. For improving the transparency, quality and legitimacy of grant allocation practices, it would therefore be important to uncover the details of the de facto (implicit and explicit) applied criteria. Which characteristics of applicants do panel reviewers value the most and how do they reach agreement within panels?

## 2.3 Theoretical background

The word 'talent' clearly has a positive connotation, but there is no general consensus on the exact meaning of it. A highly debated issue, for example, is the origin of talent: is talent innate or

acquired (e.g. Baron-Cohen, 1998; versus Howe, Davidson & Sloboda, 1998)? A recent quantitative study on career grant allocation showed that except for a few positive outliers (top talents), no evident pool of talents could be identified based on review scores (Van den Besselaar & Van Arensbergen, 2013). Thomas and Nedeva (2013) recently developed a multidimensional framework of 23 elements to characterize talented researchers, based on an extensive literature review. Examples of these elements are geographic and workplace mobility, demographic variables and the amount of academic service tasks undertaken.

### 2.3.1 Symbolic capital

In this paper we relate the notion of talent to 'symbolic capital' (Bourdieu 1986). Bourdieu discerned several forms of capital, most importantly economic (financial resources), cultural (education and upbringing) and social (relations and networks) capital. In every field, like politics, arts, economy and science, there is competition for accumulating as much capital as possible. The general distribution of these types of capital and the way they can be accumulated is field specific. The various forms of capital can be converted into other forms, most importantly into symbolic capital. Symbolic capital concerns the reputation and prestige accredited to someone, based on the recognition of his accumulated capital. In science this is mainly determined by the judgment of peers. Evaluation processes - whether it concerns reviewing scientific manuscripts for publication or assessing academics for recruitment or promotion practices - can be charac-terized as self-governing as they are mainly executed by academics themselves. Academic reputation and quality therefore lives by peer recognition. We relate symbolic capital to talent selection, since assigning the prestigious label of talent can be seen as recognition of a person's accumulated capital.

What is considered to be valuable capital and how it can be accumulated, is determined by the academic *habitus*. According to Bourdieu's concept of *habitus*, academics generally have internalized certain dispositions and norms in response to objective conditions they encountered through their academic work. The habitus is shaped by past experiences and guides current behavior and thinking within a specific field. It refers to the embodiment of certain dispositions, so not only at the explicit conscious level (Bourdieu, 1988). From the perspective of talent selection, habitus determines what is considered to be valuable capital, as it relates to dispositions academics should have or should develop in time. People are not only assessed on their current dispositions, but also on their potential, what are they expected to become, reflecting the habitus of the more established academics. Looking into how academics decide who is talented and who is not, will provide more insight in which qualities of academics the scientific community value the most and are part of the academic habitus.

In this study we modify Bourdieu's forms of capital and differentiate between professional, individual and social capital, following the example of Van den Brink (2009), who applied the concept of symbolic capital to appointment practices within academia. Professional capital involves skills, experience and achievements related to research, teaching and management. This is mainly assessed from track records, e.g. in terms of years of experience, former employers, number of publications and acquired funding. Formal assessment criteria predominantly reflect professional capital. Individual capital has more of a subjective character and is about personal traits and motivation. For example creativity, perseverance, commitment and likeability relate to individual capital. Finally, social capital is defined as consisting "*of an aggregation of networks*

*and these networks provide access to certain resources and positions of power*" (Van den Brink, 2009: p. 145). Social capital is not only valuable as it itself can be converted into symbolic capital, it also serves as an accelerator for turning accumulated professional and individual capital into symbolic capital: academic prestige. As science is turning more into a social activity, and as collaboration is of growing importance, we chose to extend the concept of social capital with those skills and traits needed for interaction and collaboration - the skills needed for creating and maintaining social networks. Thus, some of the skills that generally are considered as professional or individual capital, (e.g. communication skills, ability to collaborate, and social attitudes), are treated in this study as (conditions for) social capital.

To summarize, in this paper we study symbolic capital in the context of talent selection, as the recognition of professional, individual and social capital. We will identify the skills and traits characterizing talent and categorize them as one of the three types of capital (see table 3).

### 2.3.2 Academic talent selection

An inventory of definitions of talent in policy documents of several Dutch universities (Thunnissen, Fruytier & Van den Brink, 2010), shows that the general descriptions of talent leave considerable room for interpretation: *talents are people who perform better than expected based on their age and/or experience* (p. 20), or *talents are they of whom is expected to be able to shortly acquire a position as an associate professor* (p. 19). The main criteria for assessing young talents are publications, study and promotion results, honors degrees, awards, grants, and international experience. These criteria primarily relate to professional capital. To a lesser extent individual capital is mentioned within policy documents, such as motivation and drive.

Evaluation of scientific quality is often carried out in panels. To understand how talent is evaluated and selected in these panels, it is not enough to only study characteristics of the talents and the reviewers. Panel decisions are influenced by, and the result of, group interaction, making group and context characteristics important variables to include. Luukkonen (2012) studied review processes of ERC panels and showed how panel decisions are steered by customary interpretative and deliberation rules. From literature reviews on this type of panel reviewing, we learn that, for example group composition, group dynamics (e.g. discussion, sharing of information, power relations), characteristics of the procedure and contextual factors (e.g. budget, time pressure, accountability) can strongly affect the decision outcomes (Olbrecht & Bornmann, 2010; Van Arensbergen, Van der Weijden & Van den Besselaar, *forthcoming*). These factors impede the transparency and predictability of decision-making processes. However, as this type of panel evaluation involves interactions between human beings, it needs to be considered as a social and emotional process. Therefore, it is impossible to completely rule out any form of subjectivity (Lamont, 2009).

In this paper, we approach the process of talent selection as a strongly subjective process. We study it as fully as possible by investigating the decision-making process on both the individual and group level. We look at how reviewers use formal procedures and interpret formal criteria in their own way when evaluating grant applications. Furthermore, we analyze panel discussions and the way panels reach their final allocation decisions. By studying the process of personal grant allocation, we aim to get a better understanding of how symbolic capital is ascribed within academic talent selection.

## 2.4 Data and methods

Data for this article consists out of 29 semi-structured interviews with members of grant panels. All the respondents were involved in reviewing and allocating personal career grants in 2009 for two funding programs within the Talent Program called Innovational Research Incentives Scheme (*Vernieuwingsimpuls*) of the Dutch Research Council (NWO). Table 1 gives an overview of these funding programs, the Early Career Grant (ECG) and Intermediate Career Grant (ICG) scheme. This grant scheme is not limited to a certain scientific domain, but has eight domain panels and one interdisciplinary panel. In this paper we grouped the eight domains into two main domains: Social Sciences and Humanities (SSH), and Natural and Technical Sciences (NTS).

**Table 1** Innovational Research Incentives Scheme

| | | Early career grant (Veni) | Intermediate career grant (Vidi) |
|---|---|---|---|
| Career conditions | | 0-3 years after PhD | 0-8 years after PhD |
| Funding | | Max. €250.000 per grant | Max. €800.000 per grant |
| Duration | | 3 years | 5 years |
| Number of grants | | 150 per year | 85 per year |
| Acceptance rates (%) | 2002 | 25 | 17 |
| | 2005 | 22 | 26 |
| | 2008 | 18 | 21 |
| | 2012 | 16 | 14 |
| | 2013 | 15 | 20 |

Source: www.nwo.nl/vi

Rathenau Instituut

From the list of about 220 panellists active in 2009, we invited 40 of them across the various domains for an interview. Efforts were made to attain comparable numbers of women and men. The large majority was willing to participate: five people did not reply and seven were willing but unable to participate due to time constraints or illness. The respondents are predominantly associate or full professor and come from various scientific domains, from social sciences to life sciences (see table 2 for more details). Most of the respondents have been involved in this type of grant allocation for several years, have experience in internal selection processes at the university, and have been active as (national and international) peer reviewers.

**Table 2** Overview of the respondents per program, domain and gender

| | ECG | | ICG | | |
|---|---|---|---|---|---|
| | Male | Female | Male | Female | Total |
| Social Sciences & Humanities | 4 | 5 | 2 | 3 | 14 |
| Natural & Technical sciences | 7 | 3 | 2 | 2 | 14 |
| Total | 11 | 8 | 5* | 5 | 29* |

* including one panellist for cross disciplinary applications which is not categorized as SSH or NTS.

Rathenau Instituut

The research fieldwork was conducted in the beginning of 2012. Most of the interviews were conducted at the respondent's workplace (27), two were done by telephone, with an average duration of one hour. Respondents were asked about identification of talent in their daily academic work (e.g. describe young academics that clearly stood out in their group or who they really would have liked to retain). As this type of identification concerns the respondent's general view on talent apart from any specific selection or appraisal procedures, we refer to this as general talent identification. We feel this provides a useful frame of reference to better understand specific types of talent selection. Then we asked about talent selection within the grant panel (e.g. how do they review the grant applications; what criteria do they use in the different phases of the selection process; how do they recognize the top talents; how are the applicants discussed within the panel; how does the panel reach the final allocation decisions?). This we will call concrete grant talent selection, because of the presence of concrete selection procedures and because actual selection takes place.

All interviews were audiotape recorded, transcribed verbatim, send back to interviewees for authorization and coded using the software program Atlas.ti (see table 3 for summary of this code scheme).

**Table 3** Most important codes describing the three types of capital

| Professional | Individual | Social |
|---|---|---|
| Awards | Ability to work hard | Ability to motivate others |
| Broad expertise | Ability to work independently | Being proactive |
| Clear presentation (interview) | Ambition | Being social |
| Comprehensible proposal | Authenticity | Communication skills |
| Cum laudes | Enthusiasm | Fit in a group |
| Elaborate proposal | Goal directed | Having a large network |
| Grants | Leadership skills | Persuasiveness |
| International experience | Originality | Social skills |
| Previous employers/institutes | Perseverance | Team spirit |
| Publication record | Self-consciousness | |
| Writing skills | Willingness to learn | |



## 2.5 Results

First we focus on how talent is generally identified in daily work at the university, before shifting to talent selection in the concrete context of grant allocation. At the end of this paper we will look into the process of group decision making and which difficulties panellists face in the final phase of the selection process. As our main focus is not on gender, career stage or disciplinary differences, we will not structurally compare men and women, nor the ECG and ICG schemes, nor the NTS and SSH domains when describing our results, but only where we found important differences.

### 2.5.1 The general concept of scholarly talent

Real talents were said to be easily and widely recognized. They are those who do not excel on one single dimension, but on various dimensions, combining all forms of capital. Talent is multi-dimensional. Related to professional capital talented academics distinguish themselves from others by having broad expertise, excellent writing skills and above all a high productivity. Productivity is positively assessed when one has a high number of publications, also first authored, preferably in prestigious journals with a high impact factor. This is in line with the formal criteria formulated in many career policy documents.

Although young academics need to have accumulated considerable professional capital to be identified as talent, respondents emphasized talents above all stand out from the majority because of their individual capital. Especially female respondents strongly value this type of capital as more than half of all traits mentioned by women had to do with personality and motivation. The personality trait that is central to the academic habitus and mentioned most frequently - by both men and women - is the ability to work very hard. People who lack this ability will never make it in science, especially in top science. This also means young researchers need to be very ambitious and eager, possess a strong drive to become a top researcher, and work with great enthusiasm and passion. Since the academic world is a competitive world and set-backs are unavoidable, perseverance is considered to be an exigency, a crucial part of the academic habitus. Top talents do not let themselves be demotivated by rejections and negative reviews, but they learn from them and use these experiences to grow. Furthermore, talents are willing to learn, and are strongly self-conscious and goal directed. A last personal trait, which seems to be more important within NTS compared to SSH, is the ability to work independently and 'think for yourself'.

Being a social person is found to be a primary requirement in order to be considered a talent. As science is not purely an individual activity, but becomes to an increasing extent a social activity, it is found of great importance to fit well in a group and to have good social skills. Social skills seem to become a more important part of the academic habitus over time. In line with Bourdieu's meaning of social capital, young talents are described as people who already created their own significant network and are not just embedded in their professor's network.[2] Having a strong network is seen as meaningful in two ways. It shows they are good academics, as they are acknowledged by other people who they are connected with. But it is also considered a strategy in order to increase your career opportunities and academic success. If you really want to be successful in science you need to have the right connections and the ability to make and maintain these connections. Interestingly, having your own network as characteristic of talent was only mentioned by women, not by men. This could be explained by earlier research (Bleijenbergh, Benschop & Vennix, 2013; Van den Brink, 2009) suggesting that networks play an important role in distribution of information, resources and career opportunities. As men are overrepresented in academia and have more homogenous networks than women (Brass, Joseph, Greve et al., 2004), women may be disadvantaged and therefore more aware of the importance of having a good network.

2 For an indicator based on similar ideas: Van den Besselaar et al. (2012).

Other social characteristics valued in talents are having a strong team spirit, being interested in others, willing to help others and being able to motivate them. Finally, talents have good communication skills, take a lot of initiative and they are proactive.

> *"On the one hand you want to hire the best people, but they should also fit in the group. Some people are very good, but completely antisocial. […] we are all just human beings and we want someone in our group who is a nice person. There are plenty of people with whom in theory I could collaborate with perfectly content wise. But the chemistry is not right or there are little things of which you think, am I going to invest my time in this?"* (male, interdisciplinary)

To sum up, the majority of characteristics of talent valued in general daily academic work can be described as features of individual capital: personal traits, motivation and ambition. For women this holds even stronger than for men.

However, one type of professional capital was found to be of special importance: acquired grants. A large majority of the respondents do not consider these simply as part of track records, but as a significant indicator in itself of previous recognition of talent. In times of severe competition and generally low allocation rates, grants have an important value. They provide the receiver with both financial resources and prestige. Especially personal career grants as the ERC grants and the Dutch Veni- and Vidi-grants are considered very meaningful for young researchers.

> *"Currently it is extremely important to enter the Veni- and Vidi- trajectory or something similar at a very early stage. From the beginning of your career you should have received some sort of mark indicating you are on the right track. Those will get a plus for sure, they successfully passed the procedure. And the selections are very heavy, everyone knows. There you don't have a 80% chance, but more likely the opposite of 10 to 20%. So if you have survived, it gives you the mark 'top', meaning you are obviously very good."* (male, NTS)

As the selection is tough, those who managed to receive a grant are assumed to be excellent researchers. In the sociology of science this phenomenon is also known as the Matthew effect. Scientific credits like awards and grants are more often allocated to researchers who received them before, than to those who haven't, even when the quality of their work is similar (Merton, 1968). In terms of conversion of professional capital (grants) into symbolic capital (prestige), grants have a very high exchange rate. As these personal career grants are seen as important indicators of talent - also within career policy - we will in the next section take a closer look at this process of grant allocation.

### 2.5.2 The concrete concept of talent within grant allocation

The process of grant allocation within the Innovational Research Incentives Scheme (*Vernieuwingsimpuls*) consists of several evaluation phases.3 The criteria used to evaluate the applications are found to vary across the several evaluation phases. For this reason we successively describe the various phases and which criteria are primarily used by individual reviewers in

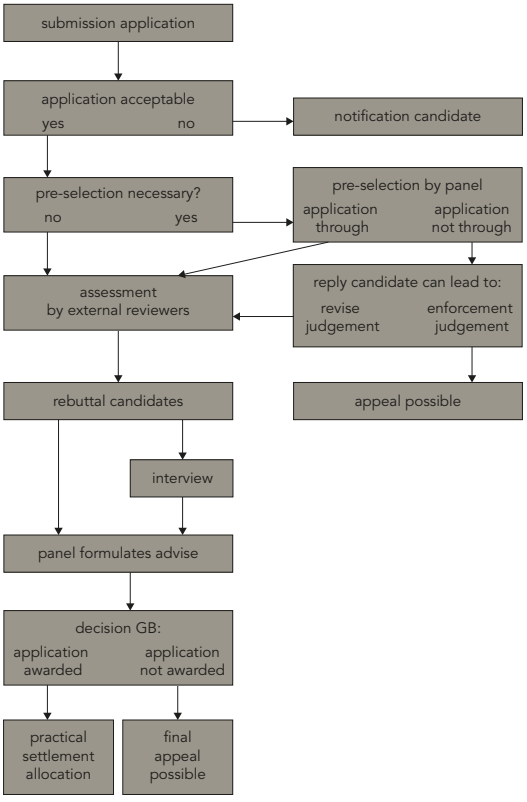
3 See Van Arensbergen & Van den Besselaar (2012) for more details on the allocation procedure.

each phase. After subsequently comparing this concrete case to the just described general identification of talent on the individual level, we will describe how talent is deliberated within panels and how panels reach agreement with regard to talent selection.

**Pre-selection: many applications, limited information**

In case of a much higher number of applications compared to number of grants, the allocation procedure involves pre-selection, as often occurs. For the pre-selection panellists receive the research proposals and the curriculum vitae of the applicants. Applications are not yet sent to external reviewers, who can be considered to be the experts in the specific research area. Panels are composed in such a way that their members cover a broad expertise. This implies that the panel as a whole is assumed to have the expertise to review all proposals as good as possible. At the same time this implies that a single panellist lacks knowledge to accurately review all proposals. Respondents indicate to be aware of this and many of them consider the broadness of the set of applications as problematic and impeding the selection process. Especially taking into account that the rejection rate in this phase is often about forty percent.

**Figure 1** Schematic overview selection procedure Innovational Research Incentives Scheme (Veni/Vidi)



> *"You need to review and judge research and performance of people from a completely different research field. And within the social sciences these differences are really huge. Certainly last time, we had people from psychology to law and so on. And yes, the tradition of publishing and way of working, everything is so different that you really are comparing apples and oranges. […] At the same time it is an enrichment as you are confronted with things you consider to be normal while others don't, and the other way around. But it is quite difficult."* (female, SSH)

Because of the heavy workload and the wide range of topics, panellists tend to primarily focus on the curriculum vitae of the applicant. The curriculum vitae, proof of mainly professional capital, is found to be generally easy to review for all applicants, regardless of disciplinary proximity.

> *"The applicant can be evaluated rather objectively. You can put a list next to it and count like, how many publications, how many awards, grants, honor degrees, that kind of things."* (female, SSH)

List of publications is mentioned as the most important indicator of talent in this first selection phase. Secondly, especially according to female respondents, there should be a notification of international experience. More generally, respondents use the institute, lab or organization itself an applicant previously worked at as a quality criteria. However, this only holds for the NTS domain, as in the SSH domain it was only mentioned once. Similar to what we found with regard to general talent selection, honor degrees, awards and grants are highly valued as professional capital.

In this first evaluation phase the emphasis is on professional capital, as three quarters of all criteria mentioned are related to this type of capital. To a much smaller extent panellists review individual and social capital. These types of capital are found harder to assess based on résumé's and written proposals. Nonetheless panellists do look for indications of authenticity, independence, and leadership skills (individual capital) and for quality of the applicant's network (social capital). They mainly ground their assessment on (co-)authorship patterns. People are positively evaluated when they published without their promoter, with various researchers from other (international) institutes, and on topics different from their PhD research. They are seen as having clear goals and ideas about what they personally want to accomplish, as being able to create their own niche and network (authenticity), and to work independently.

Although the applicant's résumé is found to be most decisive, abstracts of research proposals are important too. All panellists, regardless of specific expertise, should be able to generally understand the abstract. If it is too specific or too much jargon is used, it will not be evaluated positively. Applicants are expected to be able to inform a broad audience about their research ideas.

> *"When I started, I remember in the first round I returned five or so of the thirty applications I had to review. I said I was not able to assess them. They were far out of my working range. In the second year, you don't do that anymore, as for the third year. Then you may think, oh no, not another one you don't understand. But after all, it is up to the applicants to make clear to everyone who reads the abstract what they aim to do. If they succeed it is fine, it is a good application. If they fail, it is not a good application."* (female, NTS)

In short, in the pre-selection phase the emphasis clearly is on professional capital as assessed from résumés. Besides having an impressive publication list, including first author publications, independent from the promoter and on a different topic, talents are able to communicate their research ideas in a clear and generally comprehensive way.

**Involving external peer review: bringing in more expertise**
Next, panels have to select applicants for the interview round. Proposals are reviewed more thoroughly beyond résumé and abstract. At first proposals are more globally assessed, and panellists determine their level of expertise regarding the specific topic. When one concludes to have enough expertise, (s)he will critically review the proposal in more detail. Otherwise, (s)he will leave the thorough evaluation of the actual content of the proposal up to external reviewers, who are the experts in the specific research area.

In this phase of the process applications are sent to (generally two or three) external reviewers based on their specific expertise. In an earlier quantitative study on grant allocation within the same Dutch research council and funding program, the external reviews were found to play a modest role in the selection process, which are eventually decisive (Van Arensbergen & Van den Besselaar, 2012). This can be explained because panellists indicated to not automatically take over these expert reviews, but evaluate and weigh them. They primarily want to understand how the external reviewers determined their score. When external reviews lack a clear motivation or when panellists disagree, they can decide not to take it into account when formulating their own score.

> *"Some external reviews are very good, some are very bad. It is also possible to put the external reviewer offside, I agree. We had several situations in which we thought that the review did not come up completely objectively. This person is only trying to slice down the applicant instead of giving serious feedback. You also check if your opinion on the applications matches the opinion of the external reviewers."* (female, SSH)

When panellists evaluate the quality of the research proposal the focus is predominantly on elaboration. It should contain a complete description of the research idea, context, methodology, relevance etc. No gaps or any missing information. As in the pre-selection, generally understandable phrasing is strongly valued. Other important criteria are originality and innovativeness of research topics, and feasibility of proposals. However, some proposals describe very innovative research questions, but lack any clear evidence that the applicant will be able to answer them (in time). These proposals score lower. Furthermore panellists want to be convinced that the proposal is really the applicant's work. It should be part (or the beginning) of their own line of research, not of their PhD supervisor. Just as in the pre-selection phase, authenticity is highly valued, showing their personal contribution to the proposal, and giving it their own clear signature.

To summarize, in this phase the emphasis is still on professional capital, as the most important criteria used by panellists are track record and a clearly elaborated and feasible grant proposal. Elements of individual capital are also assessed to some extent. Talents are those applicants who are creative, innovative, and most of all authentic.

**Selection after the interview: having been face to face with applicants**

In the next stage a selection of applicants is reviewed in a face-to-face interview of about thirty minutes. Applicants present their proposal and answer questions of panellists. In line with earlier research, panellists indicate this is a very decisive phase in the selection process (see also Van Arensbergen & Van den Besselaar, 2012). For some applicants it is a clear turning point, meaning a considerable shift in their ranking.

> *"I noticed that even if you received excellent external reviews, it can still go wrong completely. That is the most important role of the panel, you assume that if one has real talent, it will show in an interview of twenty minutes. You ask a few questions and then one has to think flexibly enough to respond with 'yes indeed, when this won't work out I still have this or that theory'. So it has to be someone who thinks about research in a realistic and flexible way and does not only have excellent papers. You notice rather quickly in interviews who you really consider to be a talent, not only on paper but as a real person."* (female, NTS)

First of all talents are able to give a clear presentation, with a right balance of showing academic expertise and using a generally understandable language at the same time. Comments from the external reviewers given to them in the previous selection round, need to be taken up or contested with a clear explanation. Furthermore, the way in which applicants answer the questions is strongly assessed. For panels, this serves as an opportunity to test several of the criteria applied previously to the written application, predominantly authenticity. Where they can have received strong support from e.g. their promoter while writing the proposal, reflecting on the proposal they fully have to do themselves during the interview. This gives panels more room to assess individual capital and their academic habitus, as applicants' answers better reflect personal skills and ideas.

> *"Persuasiveness and enthusiasm, but also if you don't have that, enthusiasm needs to be there [during the interview], otherwise you're out. An extra check whether the person clearly comprehends [the topic] and whether it is really his own thing and the proposal is not written by someone else. You try to figure out if the person really is into it. As for content not much happens anymore, that happened beforehand. […] I believe it are these things. They are often hard to prove, they are intuitive."* (male, NTS)

As this male respondent indicates, during the interview intuition comes into play. Writing and presenting a proposal can be improved by training, and panels can evaluate them with a certain extent of objectivity. However, especially in this phase of the evaluation process, more subjective criteria related to individual capital play an important role, mainly enthusiasm, perseverance and ambition. Because it is all about allocating personal career grants, panellists want to be convinced that applicants really want to and are able to conduct the research project. Enthusiasm needs to be accompanied by strong persuasiveness. As this has to do with communication skills, this is linked to social capital. Finally panellists look at personality of applicants. Since there are only a few grants to allocate, they prefer allocating grants to someone they like instead of someone who appears arrogant or unfriendly.

To summarize, also in this evaluation phase, the emphasis is on professional capital. Academic skills evaluated earlier in the procedure, are tested during the interview. Does the applicant really have the knowledge and skills to conduct the proposed research project? Furthermore, during the interview much more attention is paid to individual, and to a smaller extent also to social capital. Face to face contact enables panels to evaluate the real person behind the proposal. Talents are those who succeed in convincing the panel of their authenticity, enthusiasm, ambition, perseverance, and excellent communication skills. We found that for male panellists the shift in focus during the interview from professional towards individual capital is larger than for females. The same holds true for the NTS domain compared to the SSH domain.

### 2.5.3 Tension between general and concrete talent identification

After asking the respondents how they assess applicants and applications in each phase of the review process, we also asked them to describe the top talents they came across during their grant panel work. In table 4 we listed the ten characteristics that were mentioned most often by the respondents both in general (first part of study) and concrete assessments (in grant panels), with the most frequently named traits at the top.

**Table 4** Top ten characteristics of talent in general and concrete talent evaluation[a]

| General evaluation | Concrete grant evaluation |
|---|---|
| Being social (sc)[b] | Publication record (pc) |
| Acquired grants as previous recognition (pc)[c] | Elaboration research proposal (pc) |
| Ability to work hard (ic) | General comprehensiveness (pc) |
| Ambition (ic) | International experience (pc) |
| Publication record (pc) | Authenticity (ic) |
| | Enthusiasm (ic) |
| Ability to work independently (ic) | Originality (ic) |
| Enthusiasm (ic) | Self-consciousness (ic) |
| Perseverance (ic) | |
| Writing skills (pc) | Ambition (ic) |
| Being proactive (sc) | Hot topic in research proposal |

Rathenau Instituut

a Characteristics within the same cell were mentioned just as often by the respondents and are therefore ordered alphabetically.
b sc = social capital, pc = professional capital, ic = individual capital.
c Although grants can be considered professional capital, respondents referred to them more directly as symbolic capital: previous acknowledgement of prestige.

The main difference between the lists of traits is the variation in types of capital that are valued in talents. Within the top ten of general evaluation all three types of capital are represented, with number of acquired grants being a very important indicator for talent. In concrete grant evaluation the emphasis is on professional capital, complemented with several types of individual capital but no social capital.

### 2.5.4 Deliberating talent within panels

Next, we will examine the decision-making process at the group level. How is talent discussed within panels and what eventually decides who receives formal recognition as being a talent? Both before and after the interview panellists meet to discuss their evaluations and come to a final priority ranking. During these meetings each proposal is generally introduced and commented upon by one or two panellists who were instructed to prepare a more extensive review. After this other panellists have the opportunity to respond, give their opinion or ask questions.

Panel discussions offer panellists the opportunity to explain criteria specific to their discipline, e.g. publication practices or research methods. As mentioned before, many panellists conceive the broad scope of the total set of applications in terms of topics and research fields as problematic. By exchanging information with the panellists who are experts on the specific proposal, panel meetings enable them to adjust their evaluation accordingly. The same holds true for evaluating the innovativeness, feasibility and relevance of proposals. In order to accurately assess these elements, one needs to be informed of the current state of the art of the related research field.

> *"You know the journals from the field you publish in yourself. You know the status and reputation of those journals and maybe of a few from outside your field, but of many journals you don't know this. We had interesting discussions about this, in which an applicant was about to be put aside because he lacked top publications, when a panellist said: 'But this is THE journal in this field, so it won't get any better than this.'"* (male, SSH)

In general the panel agrees with and follows the opinion of the field expert(s) within the panel. Lamont (2009) described this as adhering to the rule of 'deferring to expertise'. Experts generally have affinity with the topic, which can make them put extra effort in convincing the other panellists of the strengths of that application. But their knowledge can make them more critical too, identifying more easily the weaknesses of applications. Panellists are also found to be quicker enthused by topics they do not know much about. Some fields of research seem to be more attractive for non-experts, because topics are more appealing. In these situations the experts sometimes consider it their task to temper this enthusiasm, by e.g. explaining flaws in the research design.

> *"Compared to earth sciences people are generally easily enthused [about life sciences]. So often I say 'well, this is not that interesting' instead of 'why don't you understand that this is very interesting?' But yeah, this certainly does not hold for earth sciences. They often had to convince us that it was very interesting. Those times I didn't see this."* (male, NTS)

With regard to top talents, the majority of the respondents indicated the entire panel easily recognized them and not much discussion is needed to determine the top three.[4] Also the least

4 This could not be confirmed in an earlier study on talent selection in grant panels (Van den Besselaar & Van Arensbergen, 2013). In most panels no clear top could be discerned based on panel review scores, especially not in the beginning of the selection process.

impressive applicants are said to be easily identified. Most deliberation time is spent on the large middle area in between. Quality differences within this group are very small, leading to a rather arbitrary boundary between just selected and just rejected applicants (see Van Arensbergen & Van den Besselaar, 2012).

> *"Actually only the top 2 is evident. Of numbers 3, 4, 5 and so on, I have the feeling of, yeah, it could have been otherwise. It is all being exactly calculated. You have to assign scores including decimals. This gives a kind of impression of quasi exactness, but it isn't."* (female, SSH)

The average scores within this group are almost the same, but applicants generally vary on different criteria. One may have a better track record, but less international experience, or one gave a more convincing presentation but has a less innovative research proposal. Therefore it is very important which criteria the panellists emphasize during the discussion. This is found to be dependent on various factors, like panel composition and the comments made by the first speaker (see also Lamont, 2009; Langfeldt, 2002; Luukkonen, 2012). The composition determines the available expertise, but also more subjectively the affinity with the research topic or methods within the panel. What the first speaker starts the discussion with, is also found to be decisive according to the respondents. The strong or weak points of the applicant mentioned first, are strongly supported by other panellists.

> *"So when someone is assessed, what the first reviewer states will set the tone. When he completely tears down the proposal, the rest is only damage control. When he praises it with arguments, they others will say, yes this is also a way to look at it, and they will immediately consider it as something positive."* (male, interdisciplinary)

Furthermore, personal preferences and the atmosphere within the panel are found to influence final selection decisions too. Although panellists are convinced that in the end the top talents are granted, they indicate to have doubts about part of the allocation decisions. Since the quality differences are small, and random (social) factors influence decisions related to the middle group, many rejected applicants could have received grants as well. Most respondents indicate they are aware that this is inherent to the review process, and they would not know how to change this. As the process of grant allocation is conducted by human beings and involves social interaction, these subjective factors can never be fully excluded (Lamont, 2009).

### 2.5.5 Talent and gender

In the interviews we deliberately did not ask explicitly for gender differences, both with regard to panellists as to applicants. As gender in general is quite a sensitive and highly debated topic, we wanted to see if respondents would bring up the gender issue by themselves. The large majority did not. Only a few times, (mostly female) respondents mentioned gender differences explicitly as in female candidates being more introvert and emotional, males being more rational and self-confident. All other references to gender were made to emphasize that gender was not an issue within the panel. Interestingly, these all came from female respondents. Women seem to be more conscious of the gender issue than men, as the following quote also shows:

I: "Okay, I asked all my questions. Is there anything you would like to add?"
*R: Yes, there is one thing I was really wondering whether you would ask me about. Namely to what extent gender plays a role in this whole process. Just briefly, in my mind it is a point of attention. I mean, it is something I'm very attentive to. Myself. So I mean, if things would happen of which I think, well this is not acceptable, then that would be problematic for me. But I have to say, I'm pleasantly surprised." (female, SSH)*

The only difference we found between male and female respondents is that in describing top talents, both as grant applicants and as general university employees, women ascribe an higher value to individual capital.

## 2.6 Conclusions and implications

Analyzing talent in terms of symbolic capital, this study contributes to previous research on talent assessment or evaluation practices in several ways. Firstly, it showed the dynamics of evaluation criteria in various phases of the grant allocation process, not only with regard to the final outcomes. We described how criteria relating to various types of capital and the weight assigned to them, change during the review process. Secondly, we showed the differences between what characteristics count as talent when (early career) researchers are selected by grant panels, and the more general notion of talent used by (senior) scholars.

To start with the second issue, the general description of academic talent reflects the various activities scholars are involved in, and it entails a broad variety of skills and traits, combining professional, social and predominantly individual capital. Within the concrete context of grant allocation, a much narrower definition of talent is used and selection is mainly based on professional capital. Counts of publications, awards, grants and international experience are key criteria used in the selection process, especially in the first phase. Only later in the procedure individual capital is added. Being highly productive is obviously an important part of the academic habitus. And, more specifically, in case of early career grants as discussed in this paper, being highly productive can be considered an importance entrance criterion.

As grants are distributed with specific aims - e.g. to do innovative research - it is not surprising that the notion of talent differs between this concrete and the more general forms of talent identification. However, it is important to create a better understanding of talent selection in grant panels, because of the considerable significance assigned to grants. Grants, more specifically the personal career grants, provide academics not only with financial resources to conduct research, but also with academic prestige and further career opportunities. The acquirement of career grants like those of the Dutch Innovational Research Incentives Scheme, is considered an important indicator of talent. Competition for these grants is strong, and having succeeded in obtaining one is seen as proof that the grantee belongs to the top talents. Actually, being in the top 10% talents in one's field is one of the formal evaluation criteria. Many respondents indicated the only possibility they saw for their talents to stay in academia was by obtaining external grants. Within recruitment, evaluation and promotion procedures, acquirement of external funding - and especially prestigious career grants - form an important criterion. Skills and dispositions needed to acquire grants have been included in the academic habitus.

When obtaining prestigious career grants becomes leading in recruitment, the tension with the more general notion of talent becomes important. There is the risk of de facto undervaluation of other aspects of individual and social capital, needed for core tasks of university staff as teaching, management and valorization. The growing importance of research grants could also have negative consequences for risky and creative research, as review procedures adopted by funding agencies for allocating grants tend to contain a conservative bias. As evaluations are constrained by the boundaries of current knowledge and reviewers generally highly value criteria like validity, plausibility and a strong publication record, researchers who move to a new field or who have 'wild' ideas, have less chance being funded (Heinze, Shapira, Rogers et al., 2009; Luukkonen, 2012). The question remains whether these 'other' skills are increasingly neglected, or whether these come up in the later part of e.g., grant selection processes. This brings us to the issue of changing characteristics of what counts as talent in grant selection procedures.

In the second part of our study, we studied the various stages of panel reviewing and deliberation, following up on research by Lamont (2009), Langfeldt (2002) and Luukkonen (2012). Talent proved to be assessed differently across the various phases of the selection process, and panels change the way they discuss applications and reach agreement. Within every phase there is an overall agreement on which skills and traits applicants need to have. However, panellists face two main difficulties: first, the broadness of the set of applications they have to review and second, the minimal quality differences within the main group of applicants. Because of the broadness of the applications under review, panels have to cover a wide area of research. Therefore, panellists can only be experts to some of the applications, making the panel usually a 'panel of generalists' (Luukkonen, 2012). Discussions about individual applications are dominated by the few 'real' experts, but all panellists are involved in evaluating the applications and in decision making. In cases where the expert panellists evaluate applications highly positively, these scores will be averaged with the other scores of panellists maybe less familiar with the specific research area, methods or subject. Consequently, high potentials which are not recognized as such by the majority within the panel will end up in the so-called grey area, in which average quality scores hardly vary - even if the experts did recognize them as high potentials.

The second difficulty panellists face is that within the large set of applicants, no clear differentiation can be made using the most common criteria (e.g. publications, grants, international experience). This is often clear in the final scores, which hardly differ between many of the applicants. This implies that other criteria are involved - explicitly or implicitly. These criteria cannot always be fully articulated, as they are often more subjective and related to the tacit dimension of evaluation (Bourdieu, 2004; Van den Brink, 2009). As grant decision making is a social process involving human interactions, subjectivity and random factors can never completely be excluded, they are inherent to the evaluation process. "*Judging academic excellence is a process shaped by real-world constraints*" (Lamont, 2009: p.155). We showed how panel outcomes are influenced by individuals' behavior (e.g. personal preferences), context specific characteristics (e.g. time restrictions) and group dynamics (e.g. influence opening statement). Panel decisions convert minor differences in quality into enlarged differences in recognition.

Due to these difficulties, many panellists indicate to question the final allocation decisions. Even though they are convinced the top talents are amongst the grantees, a large share of the

grantees are considered not to be any better than many of the rejected applicants. Many panellists indicate that decisions with regard to this share are liable to arbitrary factors and are partly a matter of luck. Consequently, one would expect a devaluation of the symbolic value of these grants. Paradoxically, despite acknowledging this level of subjectivity and arbitrariness, experienced researchers still ascribe high symbolic value to these grants.

Further research is needed to determine the actual value of these grants for early career researchers. Following up on studies by Bloch et al. (2014), Gerritsen et al. (2013) and Laudel and Gläser (2012), it is necessary to better understand if and how these grants enhance academic careers and moreover, if these researchers are more successful than those without such grants. Observational studies adding to the interesting work of Lamont (2009) and Langfeldt (2002) would create opportunities to further identify the exact review criteria, including the more tacit criteria, used within panels and the effects of group dynamics on the outcomes of panel review processes. By making these implicit criteria explicit, one may be able to avoid that applicants are assessed on different criteria, and may reduce the influence of subjectivity and random factors when selecting among applicants with almost equally good publication records. And, to what extent do these implicit criteria cover the dimensions of the more general concept of talent that resulted from this study?

## 2.7 References

Baron-Cohen, S. (1998) 'Superiority on the embedded figures test in autism and in normal males: evidence of an "innate talent"?' *Behavioral and Brain Sciences, 21*(3), 408-409.

Bleijenbergh, I., Benschop, Y. & Vennix, J. (2013). Group Model Building to support gender equality change in. Achterbergh, J., Benschop, Y., Hendriks, P. & Van de Ven, A. (Eds.): *Op zoek naar het andere: een liber amicorum voor Hans Doorewaard*. Boom: Meppel: 80-96.

Bloch, C., Graversen, E. K. & Pedersen, H. S. (2014). Competitive research grants and their impact on career performance. *Minerva, 52*, 77-96.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.) *Handbook of Theory and Research for the Sociology of Education* (Vol. 1st edition, p 241-258). New York: Greenwood.

Bourdieu, P. (1988). *Homo Academicus*. Stanford: Stanford University Press.

Bourdieu, P. (2004). *Science of science and reflexivity*. Cambridge: The University of Chicago and Polity Press.

Brass, D. J., Joseph, G., Greve, H. R., & Tsai, W. P. (2004). Taking stock of networks and organizations. *Academy of Management Journal, 47*(6), 795-817.

De Boer, H. & Jongbloed, B. (2010). *Analyse van universitaire jaarverslagen 2009*. Achtergrondstudie voor Rathenau Instituut. Enschede: CHEPS.

De Jonge Akademie (2010). *Rendement van talent. Aanbevelingen voor motiverend en stimulerend loopbaanbeleid*. Amsterdam: De Jonge Akademie.

Gerritsen, S., Plug, E., & Van der Wiel, K. (2013). *Up or out? How individual research grants affect academic careers in the Netherlands* (No. 249). Den Haag: CPB Netherlands Bureau for Economic Policy Analysis.

Heinze, T., Shapira, P., Rogers, J.D. & Senker, J.M. (2009). Organizational and institutional influences on creativity in scientific research. *Research Policy, 38*, 610-623.

Howe, M.J.A., Davidson, J.W. & Sloboda, J.A. (1998) 'Innate talents: reality or myth?'. *Behavioral and Brain Sciences, 21*(3), 399-407.

Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment*. Cambridge: Harvard University Press.

Langfeldt, L. (2002). *Decision-making in expert panels evaluating research. Constraints, processes and bias*. Norwegian Institute for Studies in Innovation, Research and Education, Oslo.

Laudel, G. & Gläser, J. (2012). The ERC's impact on the grantees' research and their careers. EURECIA Understanding & Assessing the impact & Outcomes of the ERC Funding Schemes. Work package 4 summary report. http://www.eurecia-erc.org/wp-content/uploads/EURECIA-ImpactOnResearchContentAndCareers-SummaryReport.pdf.

Lepori, B, Van den Besselaar, P., Dinges, M., Potı`, B., Reale, E., Slipersaeter, S., Theves, J. & Van der Meulen, B. (2007). Indicators for comparative analysis of public project funding. Concepts, implementation and evaluation. *Research Evaluation, 16*(4), 243-255.

Luukkonen, T. (2012). Conservatism and risk-taking in peer review: emerging ERC practices. Research Evaluation, 21, 48-60.

Merton, R. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science, 159*(3810), 56-63.

Neufeld, J., Huber, N. & Wegner, A. (2013). Peer review-based selection decision in individual research funding, applicants' publication strategies and performance: The case of the ERC Starting Grants. *Research Evaluation, 22*(4), 1-11.

Olbrecht, M. & Bornmann, L. (2010). 'Panel peer review of grant applications: what do we know from research in social psychology on judgment and decision-making in groups?'. *Research Evaluation, 19*(4), 293-304.

Thomas, D. & Nedeva, M. (2012). Characterizing researchers to study research funding agency impacts: the case of European Research Council's Starting Grants. *Research Evaluation, 21*, 257-269.

Thunnissen, M., Fruytier, B. & Van den Brink, M. (2010). *Beleid en beleving. Onderzoek naar jongetalentenbeleid op Nederlandse universiteiten*. Den Haag: SoFoKleS.

Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (*forthcoming*). The selection of talent as a group process. A literature review on the social dynamics of decision-making in grant panels. *Research Evaluation*.

Van Arensbergen, P. & Van den Besselaar, P. (2012). The selection of scientific talent in the allocation of research grants. *Higher Education Policy, 25*, 381-405.

Van Arensbergen, P., Hessels, L.K. & Van der Meulen, B. (2013). Talent Centraal. *Ontwikkeling en selectie van wetenschappers in Nederland*. Den Haag: Rathenau Instituut.

Van den Besselaar, P., Sandström, U. & Van der Weijden, I. (2012). The independence indicator. E. Archambault, Y. Gingras, V. Lariviere (eds.) *Science & Technology Indicators 2012*. Montreal: OST & Science Metrix, 131-141.

Van den Besselaar, P. & Van Arensbergen, P. (2013). Talent selection and funding of research. *Higher Education Policy, 26*, 421-427.

Van den Brink, M. (2009). *Behind the scenes of science. Gender practices in the recruitment and selection of professors in the Netherlands*. PhD thesis, Radboud Universiteit Nijmegen.

Van den Brink, M., Fruytier, B. & Thunnissen, M. (2012), Talent management in academia: performance systems and HRM policies. *Human Resource Management Journal, 23*(2), 180-195.

Waring, M. (2013). All in this together? HRM and the individualisation of the academic worker. *Higher Education Policy, 26*, 397-419.

# 3 The selection of talent as a group process A literature review on the social dynamics of decision-making in grant panels[1]

Talent selection within science is increasingly performed by panels, e.g. by reviewing grant or fellowship applications. Many studies from fields of sociology of science and science policy studies have been conducted to identify biases and predict outcomes of these processes, mainly focussing on characteristics of applicants, applications and reviewers. However, as panel reviewing entails social interaction, group dynamics influence these processes. By adding insights from social psychology to current knowledge on panel reviews, we are better able to identify factors affecting talent selection and funding decisions in grant panels. By opening up this so-called black box we aim to contribute to a better understanding of the dynamics of panel decision-making. This knowledge is also relevant for various stakeholders involved in grant allocation, for applicants, reviewers and policymakers, as it can be used to improve transparency, fairness and legitimation of talent selection processes.

## 3.1 Introduction

The academic market in both the United States and most European countries is a buyer's market, and has been so for quite some years, given the strong preferences of many new PhDs and postdocs for a job at the university. Researchers who are lower in the academic hierarchy hold to an increasing extent temporary positions without prospect of permanent employment (Stephan, 2012). This shift towards more temporary contracts is mainly due to an increase of the proportion of research within universities that is based on short term external funding, like project funding or individual career grants. Consequently, opportunities for especially young academics to conduct research and develop an academic career are more and more characterized by competition for funding.[2]

A rationale behind project funding is that it strengthens competition between researchers, and therefore promotes the quality of science: only the best succeed. The ability to acquire research grants is turning into a prominent criterion in processes of academic recruitment and performance evaluation (De Jonge Akademie, 2010; Van Arensbergen, Hessels & van der Meulen, 2013). Career grants are not only a way to directly distribute financial resources amongst young researchers to conduct research, it also provides them indirectly with improved career opportunities as grants are considered significant indicators of excellence or talent (Van Arensbergen, Van der Weijden & Van den Besselaar, *forthcoming*). This line of reasoning is based on the assump-

1 This chapter has been accepted for publication after minor revisions by Research Evaluation as Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (forthcoming). The selection of talent as a group process: A literature review on the social dynamics of decision-making in grant panels.

2 To give an impression of how strong the competition for individual career grants is: for the starting grants of the European Research Council the success rate in 2013 was 9% (http://erc.europa.eu/statistics). For similar type of career grants in the Netherlands - Veni and Vidi grants - success rates in 2012 were 16 and 14% respectively (NWO, 2012).

tion that grants are awarded to the best applicants. Although this obviously is what funding agencies claim, several recent studies suggest otherwise (Bornmann et al., 2010; Hornborstel et al., 2009; Melin & Danell, 2006; Van den Besselaar & Leydesdorff, 2007, 2009; Van den Besselaar, 2013). As these funding decisions have great impact on individuals' careers, but also shape the direction of academic research (Hodgson, 1997), we consider it of great importance to create a better understanding of underlying decision-making processes. Of course, uncertainty in grant decisions is unavoidable, which prevents high predictive validity. But a better understanding of the social dynamics may help research councils to improve their selection practices and to reduce the systematic bias that may emerge from social dynamics in decision-making bodies.

The main method used to make these allocation decisions is a combination of individual peer review and panel review (peers reviewing as a group). Although peer review has been studied a lot, attempts to predict the outcomes of funding allocation processes show it still largely is a black box (Cole, Cole & Simon, 1981; Hartmann & Neihardt, 1990; Van den Besselaar & Leydesdorff, 2007).

Contributing to the unpredictability of these review outcomes is the nature of this type of decision-making: it often involves group decision-making. Often panels of peer experts are installed to assess the quality of applications and to decide on funding allocation. Reasons for installing panels mainly have to do with the size and width of the set of applications and with the weight of funding decisions. A panel of reviewers has more resources to draw on than one or two individual reviewers (information integration). And secondly, decisions made by a panel of experts (through consensus building) are considered more acceptable than individual decisions (Olbrecht & Bornmann, 2010).

## 3.2 Focus of this review

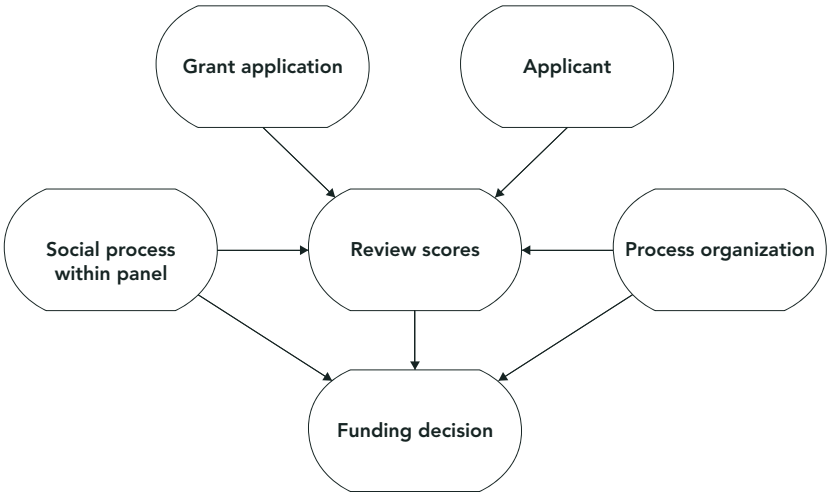
The present literature study focuses on decision-making as performed by panels, including individual peer review. Originally peer review is considered the legitimate method to evaluate scientific quality of scholarly contributions and therefore is deeply embedded in research culture. Peers are considered to be best suitable to assess scholarly quality and to distinguish inferior from meritorious research by means of critical appraisal (Hartmann & Neihardt, 1990; Langfeldt 2002). At the same time it is highly criticized as being unreliable, costly and biased (e.g. Marsh, Jayasinghe & Bond, 2008; Porter & Rossini, 1985). Most of the studies on peer review stem from sociology of science (SoS) and science policy studies (SPS).[3] They mainly deal with how review outcomes are affected by performance and characteristics of individual applicants, and by characteristics of reviewers. These studies are predominantly based on analyses of written documentation (e.g. submitted proposals, review reports and reports of meetings), interviews (e.g. with reviewers and applicants), and bibliometric data.

However, many allocation (and appointment) decisions are made in panels, which is not covered very well by peer review literature. Panel review is not the same as peer review, as panelists are often not peers. Since panels have to assess heterogeneous objects (e.g. grant applications

3 See Bornmann (2011) for a recent literature review on peer review.

covering a range of fields, research programs, job candidates), they are composed in such a way that they cover a broad range of expertise. Consequently, panelists are not all experts or peers to every object they review. Panels may also include non-scientific members, e.g. societal stakeholders, representatives from funding agencies or university board members. Furthermore, peer review generally is just part of panel review procedures. Finally, panel review is embedded within group interaction and can therefore be characterized as a social activity. For this reason we combine SoS and SPS literature on peer review with literature on group decision-making from social psychology (SP). Where the first mainly focus on how peer review affects review outcomes, the latter focusses on actual review processes. SP research predominantly deals with central mechanisms involved in decision-making processes and the context in which these are carried out. To a large extent this literature is based on experimental research. Langfeldt (2001, 2002), Lamont (2009), and Olbrecht and Bornmann (2010) also looked at social psychological research with regard to panel review. Although they describe several important mechanisms that could affect panel review (e.g. motivation losses and group polarization), based on exploratory observations of panels we know there are more factors related to panel interaction that influence allocation decisions.

**Figure 1** Basic model of grant allocation process




Rathenau Instituut

## 3.3 Methodology

The literature exploited in this study mainly comes from Web of Knowledge searches, added with Google Scholar hits. Figure 1 depicts the straightforward model we used for our literature search and to structure our review. We searched for literature using as main key words 'peer review', 'grant allocation', 'group decision-making', 'group interaction', and 'intragroup behaviour'. Searches resulted in a broad scope of literature in terms of type of research (e.g. interview studies, bibliometric analyses, historical case study analyses, (lab) experiments) and in terms of

potential factors influencing panel reviews. Results were refined based on six exploratory observations of panel meetings in 2010 and 2012, in which grant applications were reviewed and ranked in pre-selection and selection phases. We observed several issues related to group dynamics, which seemed to influence panel processes. For example, panelists varied in their motivation and contribution to panel deliberation, and similar types of information (e.g. anecdotal or shared) were not always considered evenly important. Factors identified in our observations and included in our SP literature review are social status and identity, group norms and cohesiveness, information distribution, motivation and interests, and procedural factors.

Next, we describe how characteristics of people or proposals under review affect review outcomes. For this we primarily draw upon SoS and SPS literature. Second, review processes as a social interaction between various panelists are explored in more detail. Characteristics of panels and dynamics inherent to group decision-making are further explained predominantly using SP literature. Finally, also based on SP literature we look at influences of external factors related to the organizational context in which the review process is carried out.

## 3.4 Panel review of grant applications

### 3.4.1 Explicit quality related criteria

Since funding organizations claim to fund only excellent research and the best researchers, one expects in accordance with Merton's (1973 [1942]) norms of universalism the scholarly quality of grant proposals and of applicants to be central criteria for a proposal's acceptance or rejection. However, already thirty years ago the study of Cole et al. (1981) on funding decisions within the National Science Foundation (NSF) could not confirm this. They did not find a strong correlation between funding success and *past performance* of researchers. More recent studies using different types of data sources show inconsistent results. For example, in order to build statistical discriminatory models that can replicate peer review outcomes, Cañibano, Otamendi and Andújar (2009) used curricula vitae of applicants to a Spanish research program. They found *research productivity* to be the main determinant of grant success. Other studies comparing past performance of granted applicants with unsuccessful applicants generally found that the former have higher average performance than the latter (e.g., Bornmann & Daniel, 2006; Van Leeuwen & Moed, 2012). However, as competition has become harsh, successful groups are much smaller than rejected groups, which include also many low performing applicants. This has been subject to further investigation, and researchers have started to compare successful applicants with more restricted sets of good performing rejected ones - something that changed the outcomes: successful applicants do not outperform about equally large groups of best performing applicants (Melin & Danell, 2006; Van den Besselaar & Leydesdorff, 2007, 2009; Hornbostel et al., 2009). More generally, in terms of past performance, selection processes are characterized by large numbers of false positives (granted applicants performing less than rejected applicants) and false negatives (rejected applicants performing higher than granted applicants). Bornmann et al, (2008) found percentages between 26 and 48% in two grant programs within life sciences, and Bornmann et al (2010) similar percentages in grant programs in the life sciences and the social sciences.

More recently, the focus has shifted from past performance to post performance analyses: do the selected applicants indeed prove to be better in the years after having received grants? Here

similar patterns of results are emerging - granted applicants have in average a better post-performance than all rejected (Bornmann et al., 2008) but not if compared with the best rejected (Melin & Danell, 2006; Van den Besselaar, 2013).

### 3.4.2 Implicit quality related criteria

Laudel (2006) disproves the dominant 'quality myth' in her interview study with German and Australian scientists about their research trails and funding sources. She suggested several non-quality related factors influencing funding decisions. For example, continuity of research trails, countries' general investment in research and scientists' *research field*. A research field bias was also found in several other studies. Bornmann and Daniel (2005) demonstrate that success rates for doctoral fellowship applicants working in the field of chemistry are approximately half as high as for applicants working in other fields within life sciences. The analysis of applications for post-doctoral fellowships however, does not confirm this bias, suggesting we cannot speak of an unambiguous relation between research field of applicants and grant success. A relevant issue here is whether choice of field is a 'non-quality related factor', as there are more promising and less promising research topics and fields, and selection of topics may be seen as a quality of the researcher at stake.

An important variable that should be taken into account is the research field of reviewers. In case of a *disciplinary match* between applicant and reviewer, review scores are significantly higher than when there is no match (Porter & Rossini, 1985). This can be explained in terms of *cognitive particularism*, meaning that people make decisions based on cognitive similarity, their membership in a particular scientific school of thought. "It is not that panel members are not of goodwill but that they simply do not fight so hard for subjects that are not close to their hearts" (Travis & Collins, 1991, p. 336). Consequently, proposals on topics that are unrelated to the panel members' interests may be disadvantaged here, and that may also hold for interdisciplinary research proposals. Since interdisciplinary research can be seen as a novel way of integrating expertise, real peers may be hard to identify[4]. However, research on both peer review and bibliometric assessments found no significant bias with respect to interdisciplinarity (Rinia et al., 2001).

*Status* also plays an eminent role in evaluation processes. This relates to academic status of applicants and status of their department, university or institute. Applicants with a higher academic and/or departmental status have better chances on securing grants than applicants with relatively lower status (Bazely, 1998; Bornmann & Daniel, 2005; Cole et al., 1981; Viner, Powell & Green, 2004). This shows that not only characteristics of applicants themselves are influential, but also those of institutes they are affiliated with. Another influential type of *affiliation* involves panelist affiliation. Wennerås and Wold (1997) found that higher competence scores are given to applicants who are affiliated with a panelist than to applicants without such ties. This affiliation may explain partly why the academic status of the applicants' institution plays a role: the panelists themselves may predominantly come from the same high reputation institutes. Important to emphasize is that the affiliated panelists themselves are not allowed to participate in scoring of the specific proposals. 'Neutral' reviewers seem to compensate for the absence of

4 One may use bibliometric methods to find those peers.

scores by 'biased' reviewers by raising their scores assigned to applicants associated with one of their peers. Furthermore, applicants who are themselves member of a peer review cadre have more chance to be allocated grants than applicants who lack this type of membership (Viner et al., 2004), and this is not explained by performance differences (Van den Besselaar, 2012).

A highly contested variable in peer review literature is *gender*. Related to funding decisions it was demonstrated that women receive relatively fewer grants than men (Bornmann, Mutz & Daniel, 2007). However, there consists general disagreement over the impact of gender on outcomes of peer review and grant allocation. In their well-known study Wennerås and Wold (1997) looked at applications submitted to the Swedish Medical Research Council. They observed that peer reviewers assigned lower scores to female than to male applicants, while their levels of scientific productivity were about the same. A similar study on grant applications in the Netherlands confirmed that gender matters (Brouns, 2000). How it matters was found to vary between disciplines. Whereas in some disciplines in case of equal average publication scores more men than women were evaluated as excellent, less productive women also obtained grants in others. This implies the use of double standards. Women have to perform to higher levels to be considered as qualified as men, according to both men and women (Foschi, 2004; Van den Brink, 2009). However, in accordance with several other studies (Bazely, 1998; Jayasinghe, Marsh & Bond, 2001; Marsh et al., 2008; Mutz, Bornmann & Daniel, 2012; Sandstrom & Hallsten, 2008) Ceci and Williams (2011) in their recent review on discrimination against women in science, found no evidence supporting current discrimination of women in grant allocation.[5] However, an extension and reanalysis of previous data by Marsh et al. (2009) shows it is important to distinguish between types of applications. Whereas there were no gender differences with regard to grant applications, there were differences in favor of men with regard to fellowship applications.

Decreasing gender disparities can be the effect of changed (council) policies, as suggested by several studies (Sandstrom & Hallsten, 2008; Van den Besselaar & Leydesdorff, 2007, 2009). Among scientists themselves - male and female - it is even not seen as main concern in grant assessments (Van der Weijden & Calero Medina, 2014). However, as within science in general gender disparities persist (e.g. Lariviere et al., 2013; Ranga, Gupta & Etzkowitz, 2012) this issue still needs further study and attention.[6]

As we already saw with regard to research field and affiliation, review outcomes do not solely depend on characteristics of candidates under review. Evaluation outcomes are determined by interaction between characteristics of reviewers and the reviewed. With regard to panel review there is another type of interaction significantly affecting the review outcomes: interaction between panelists. Therefore we will now take a closer look at panels and describe review processes as social interaction between panelists. We will describe various factors inherent to

5 Many studies included in their review did not use data on the performance of applicants - a general weakness of many studies on gender bias in grant decisions.

6 For example, the European Research Council recently launched the projects ERCAREER (Capturing gendered career paths of ERC grantees and applicants) and GendERC (Gendered dimensions in ERC grant selection), and the European Committee granted EGERA (Effective Gender Equality in Research and the Academia) and GARCIA (Gendering the Academy and Research: combating Career Instability and Asymmetries).

social interaction, that influence decision-making processes and are subsequently expected to affect review outcomes. In the following paragraphs we will identify factors which need to be studied in more detail, in order to determine how they affect outcomes of grant allocation processes.

## 3.5 Peer review as social interaction

As processes of grant allocation generally involve quality assessment by panels, they can be considered to be social, emotional and interactional processes (Lamont, 2009). Panel decisions are the outcome of and are influenced by group interaction. Differences in for example status and expertise of the panelists can play an important role in this type of interaction. Furthermore, group interaction can make group members motivate each other and increase the amount of information that is collected and discussed, compared to individual decision-making. On the other hand, group interaction can result in poorer decision-making, because shared responsibility creates a situation in which everyone withdraws and no one really endeavors, better known as *social loafing* (Levi, 2007). It can also encourage members to focus primarily on reaching consensus, so they are not really motivated to detect possible weaknesses in their decisions and to realistically appraise alternative decisions. This social psychological phenomenon is better known as *groupthink* (Janis, 1982). We will therefore look in more detail at panel review as a social interaction process. We will describe how specific characteristics related to the social nature of this process can affect panel decisions. Based on our observations we will successively focus on the composition of the panel, group norms and cohesiveness, information distribution, and finally we will look at the motivation and interests of panelists.

### 3.5.1 Panel composition

Several studies showed that outcomes of reviewing decisions to a great extent depend on who the reviewers are and how the panel is composed (e.g. Lamont, 2009; Van Arensbergen et al., *forthcoming*). According to Van den Brink (2009) in the Netherlands more women in appointment committees lead to higher numbers of women being appointed as full professor. The same was found in a Spanish study on promotion decisions: adding a female evaluator to the committee, increases the number of females promoted to full professor by 14% (Zinovyeva & Bagues, 2010). This indicates preferences for same-sex candidates. However, this type of bias was not found in promotion decisions for associate professors. Moreover, female associate professors were found to discriminate against female candidates from the same institution, possibly for strategic reasons.

In general scholars are asked for grant panels based on their disciplinary expertise and research experience. Often applicants may enclose to their proposal names of some reviewers they definitely do not want to be part of the panel. In some cases applicants also have the opportunity to *nominate* people for panel membership. Reviewers nominated by applicants are found to systematically give higher scores to all proposals than reviewers who are appointed by the board or otherwise (Jayasinghe et al., 2001; Marsh et al., 2008).

Another aspect of panel composition is the difference in *expertise represented* by panelists. The set of applications generally covers a broad range of topics, sometimes even from various disciplines. Consequently experts from different disciplines have to be included in the panel to enable a fair and comprehensive evaluation of all proposals. But also within a disciplinary panel, people can be considered experts on different topics or research areas. It is important to pay

attention to panel composition, since the composition sets the potential for interaction and conflict among its members. Overlap in competences is associated with better cooperation and with open conflict between scientific experts (Langfeldt, 2002). Research on decision-making also shows that groups with heterogeneous members with complementary skills take better group decisions than homogenous groups (Levi, 2007). However, the advantage of heterogeneous groups does not arise directly from the broad range of knowledge that is present in groups. Members have to be *conscious* of differences in areas of expertise. An experiment conducted by Bonner, Baumann and Dalal (2002) showed that when group members know who the experts are in reference to a specific task, they will adjust their group decision to the decision of the experts. This of course, does not necessarily mean that the group decision will be of better quality when decisions of experts are taken over. But it can be considered a stimulus for information sharing. A social psychological experiment using the *hidden profile*[7] task demonstrates that when people know who knows what, distributed information is mentioned more often and the hidden profile is solved more often (Stasser, Stewart & Wittenbaum, 1995). Hence group decision-making can benefit from diverse panel compositions as long as this diversity is evident to everyone within the panel. We will come back to the issue of information sharing later.

Furthermore, within panels there may be differences in status. By this we mean the *status* as *perceived* and *implicitly assigned* to them by other panelists. Some people might be considered to be hotshots with very good reputation and hence have a high status. Others might be seen as newcomers or relatively insignificant in their field. These perceived status differences cause unequal *power* distribution amongst group members, which subsequently will disturb communication within groups. In general, high-status members talk more and receive more attention from other members. Low-status members generally talk less or even do not talk at all when their opinions deviate from those of high-status members. This can harm decision-making processes, since not all true opinions are expressed and high-status people will not be contradicted often. Communication plays an important role in processes of group decision-making. In order for a group to perform well, it is desirable that group members trust each other and that there is open communication between them. This can be facilitated by good social relations within the group (Levi, 2007).

Finally, panel composition affects the way individual panelists identify themselves. Individuals do not have one fixed identity, but depending on the social context they are in, different identities can be addressed. Interaction between characteristics of individuals and of the specific situation determines which particular identity is activated. This process of *social identity formation* comprises two important activities, namely social comparison and self-categorization in terms of membership in particular groups (Stets & Burke, 2000). By means of self-categorization in-groups and out-groups are created, which leads to accentuation of perceived similarities between the

7 This contains a group decision-making task in which the best solution cannot be detected by individual members based only on the information they received prior to discussion. There is a difference in the information individuals have at their disposal. Prior to the group discussion, partial information is given to all group members (shared information), whereas other pieces of information are known to some but unknown to other members (unshared information). Based on the available information individuals will detect different 'best' solutions. To find the only real best solution unshared information has to be pooled during the discussion.

self and other in-group members and of perceived differences between the self and out-group members (Hogg & Terry, 2000). Grant panels can be considered by its members as (one) in-group, but can also comprise several smaller groups. Panelists possibly identify themselves with some and not with other members. For example, when people share a disciplinary background, professional status or faculty membership, this can determine the in-group identity.

Van Kleef et al. (2007) studied social identification in terms of prototypicality. A group that people identify with generally comprises one or more members that can be considered to be the group prototype. "Prototypes embody all attributes that characterize groups and distinguish them from other groups, including beliefs, attitudes, feelings, and behaviours" (Hogg & Terry, 2000, p.123). In two experiments Van Kleef et al. (2007) compared differences in negotiation behavior between prototypical and peripheral group members. Prototypical group members are those who strongly match group prototypes and those who hardly match are called peripherals. They demonstrate that in case of valued group membership, peripherals within the group are more competitive and less cooperative than prototypicals. Two studies conducted by Terry and Hogg (1996) show that social identification has an effect on the intentions and actual behavior of group members. "When social identity is salient, depersonalization occurs, such that a person's feelings and actions are guided more by group prototypes and norms than by personal factors" (p. 790).

### 3.5.2 Group norms and cohesiveness

According to Lamont (2009) panel discussions are steered by informal rules, generally known by all panelists. These unwritten rules defining appropriate and inappropriate behavior in groups, are called *group norms*. Norms usually emerge unconsciously and gradually through interactions of group members, and are not necessarily made explicit or formal. Sometimes people are unable to articulate norms which they clearly use to guide their behavior. Norms can have a strong impact on behavior of group members, even stronger than externally imposed rules e.g. by supervisors or organizational practices. Examples of group norms are that panelists are expected to give each other full liberty to express opinions without reprisal, they should be oriented towards producing consensual decisions, and they should maintain collegiality (Janis, 1982; Lamont, 2009; Levi, 2007; Marques et al., 2001; Spector, 1996).

According to Levi (2007) norms enable groups to create a clear *group identity*, as they express central values of groups and prescribe what is accepted and deviant behavior within groups. This way group members can distinguish themselves from others and have a sense of who they are as a group. The other way around, norms are found to be dependent on the social identity perceived by individual panelists. As mentioned earlier, group norms have more effect on the behavior of individuals the stronger they identify themselves as being part of a social group and not merely as unique individuals (Terry & Hogg, 1996). Social identity is related to *cohesiveness* within groups. A highly cohesive group is characterized by strong interpersonal bonds holding a group together. Group cohesiveness refers to a sense of team spirit and the extent to which group members appreciate their group membership and share group goals. Conformity to norms is found to be more likely in groups that are highly cohesive (Levi, 2007; Spector, 1996). Groups characterized by high levels of cohesion are found to be better able to communicate and work together (Beal, Cohen, Burke et al., 2003). This could lead to better group outcomes. An analysis of case histories of seven corporations compared decision-making characteristics of top manage-

ment teams in successful and unsuccessful times (Peterson et al., 1998). Group cohesiveness was one of the characteristics they studied. They found that successful decision-making groups showed more cohesiveness than unsuccessful groups. On the other hand, members of cohesive groups may want to preserve the group's relationships and therefore avoid any kind of behavior considered to be harmful. This could mean that people agree to group decisions, while they actually don't agree with it individually. According to Janis (1982) strong group cohesion is one of the important antecedent conditions for groupthink; "a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members' strivings for unanimity override their motivation to realistically appraise alternative courses of action" (p. 9). Groupthink increases chances on flaws in the decision-making process, consequently leading to poorer decisions. However, Esser (1998) who reviewed two areas of groupthink research - historical case analyses and laboratory tests - poses that group cohesiveness is not a strong predictor of groupthink. Also with regard to the effect of cohesiveness on group performance or decision-making, no unambiguous relation has been found (Spector, 1996).

### 3.5.3 Distribution of information

The main advantage of panel compared to individual peer review is that there is more knowledge available as all individuals' knowledge is pooled together. During panel meetings reviewers share their expertise and inform each other about their assessments. Generally, type of information can be classified in three different ways: *shared* versus *unshared*, *preference consistent* versus *preference inconsistent*, and *instrumental* versus *non-instrumental*.

In terms of shared and unshared information, general knowledge most reviewers have about applications can be considered to be shared information, whereas additional knowledge someone has based on his specific expertise can be considered as unshared information. An experiment using the *hidden profile task* showed that groups in which all information is shared make better decisions than groups in which some group members have unique information (Schulz-Hardt et al., 2006). As grant panelists vary on level of expertise with regard to applications they have to evaluate, there will always be both unique and shared information. In general during group deliberations more attention is paid to shared than to unshared information. Consequently shared information has more impact on the final group decision (Baron, 2005; Gigone & Hastie, 1993; Tindale et al., 2001; Winquist & Larson, 1998). This tendency would inhibit the added value of experts contributing their specific knowledge that other reviewers do not have to panel review processes. However, based on the study mentioned earlier of Bonner et al. (2002) we argue that for unique information to be influential on panel decisions, the person bringing in this information should be recognized as being an expert.

Information distribution is also affected by initial opinions or preferences of panelists. In panels characterized by divergent opinions more information is put into deliberation than panels in which there is high agreement to start with. Furthermore, heterogeneity in opinions stimulates group members to spend more time on (information steered) deliberation and results in better group decision outcomes (Scholten et al., 2007; Schulz-Hardt et al., 2006). With regard to type of information that is put into discussion, we discern information that is consistent and inconsistent with one's initial preferences. Mojzisch, Grouneva and Schulz-Hardt (2010) found in their experiment on biased information evaluation, that people paid more attention to preference consistent

information than to information that conflicted with their preferences. This effect was even stronger when confirming information was introduced by the person himself than by other group members. Whether people adjust their initial preference based on new information that is contributed to the discussion is strongly influenced by *social validation*. Affirmation of preference inconsistent information by other group members raises the perceived quality of this information (Mojzisch et al, 2008). The bias of favoring preference consistent information can be explained as a tactic: people defend their initial preference and in order to convince others they mention more information that supports their preference (Wittenbaum, Hollingshead, & Botero, 2004). But it can also be the result of more unconscious processes: people consider preference consistent information as more accurate and relevant and therefore pay more attention to it (Mojzisch et al., 2010).

Finally, we discern instrumental and non-instrumental information. Information that is relevant for and ought to impact decisions is called instrumental, whereas irrelevant information that should not affect decisions is called non-instrumental. According to Bastardi and Shafir (2000) people often give non-instrumental information instrumental value without being aware of this. In order to base their final selection decisions on thorough evaluations, review panels collect as much information as possible, both instrumental and non-instrumental. Next, also newly obtained non-instrumental information is used to make decisions, as "the very act of pursuing information may lead people to endow it with instrumental value" (p.217). As the mere act of obtaining adds weight to new information, disregarding its relevance, information that is known from the start might receive less attention than new information (Bastardi & Shaffir, 1998). This implies that for example anecdotal information about applicants mentioned by panelists rather coincidently can influence review outcomes.

### 3.5.4 Motivation and interests of panelists

Panelists might differ in their motivation to engage in allocation processes. According to Merton's norm of *disinterestedness* reviewers should not have any personal, political or economical interests interfering their assessment of applications. Applications should be assessed purely on their academic merits. Nonetheless, several types of interests are conceivable to be held by panelists, like personal, departmental, university, disciplinary, gender etcetera. For example, reviewers might find it important to fund more research in his specific field or that more women get opportunities to build academic careers. Reviewers do not always have to be completely aware of these interests, they can influence their preferences in a more subtle way. The stronger individual preferences deviate from preferences of other panelists, the smaller its contribution to final panel preferences (Tindale et al., 2001). How panelists interact with each other and the extent to which they exert themselves during review processes, is influenced by their motivation. Next we will discern various types of motivation: epistemic, social and competitive motivation.

The extent to which reviewers search for missing information and process newly obtained information depend on their willingness to exert oneself to come to accurate and well informed assessments of applications. This is called *epistemic motivation*. Interaction within groups characterized by high epistemic motivation is found to be more steered by information than by preferences and is less susceptible for reasoning errors. Furthermore, these groups are more open to deviating opinions and they develop more egalitarian and participatory interaction patterns (De Dreu, Nijstad, & Van Knippenberg, 2008). A way to increase the epistemic motiva-

tion is to make reviewers accountable for review processes (De Dreu et al., 2007). We will come back to the role of accountability when describing procedural factors.

The type of information reviewers search for and process is found to be influenced by their *social motivation* (De Dreu et al., 2006; De Dreu et al., 2008).There are two types of social motivation: proself and prosocial. People with proself motivation have a strong focus on reaching personal goals and interests. On the other hand, people with prosocial motivation focus more on fairness and common goals. Considering the general task of review panels (assessing the quality of applications) panelists are not expected to be proself motivated. Panels in which members do have personal interests and primarily strive for reaching personal goals, are less likely to reach consensus than panels focusing on common goals. This can be the result of the exchange of information being distorted, as people with proself motivation tend to neglect other people's opinions. Social motivation consequently can lead to biases in information processing (De Dreu et al., 2008). Kramer, Pommerenke and Newton (1993) found that the extent to which one takes decisions based on self-interests or on the other party's interests is also affected by the salience of a shared social identity. During a decision-making task that involved negotiation, people were found to show greater concern with outcomes obtained by the other party and to have preferences for more equal outcomes, when a social identity was salient. When a distinctive personal identity instead of a social identity was salient, negotiators focused primarily on their own outcomes guided by self-interests.

The process of grant allocation involves a certain extent of negotiation, when panelists have strong preferences and try to convince each other of these. Looking at panel review processes as a type of negotiation, panelists can be ascribed *competitive motivation*; assuming that an individual's goal achievement is negatively related to goal achievements of others (Ten Velden, Beersma & De Dreu, 2011). Two types of competitive motivation can be discerned: appetitive and aversive. People with appetitive motivation focus on outdoing their counterparts, acquiring better results. Aversive competitors try to prevent their counterparts from doing better than them, they aim at avoiding worse results than their competitors. In a series of experiments Ten Velden et al. (2011) compared negotiation behavior of appetitive and aversive competitors. Their study showed that individuals with appetitive motivation were more confident that agreements would satisfy their goals, and they more easily reached agreements. Furthermore, identical pieces of information were found to have different effects on negotiation depending on motivational goals of negotiators. This shows that the same information can be used in different ways and that motivation influences the effect specific information has on negotiation processes.

We described earlier how social identity influences interaction and negotiation within groups: prototypical members are less competitive and more cooperative than peripherals (Van Kleef et al., 2007). This suggests that panelists may use different strategies or *social tactics* in processes of decision-making, e.g. consultation, pressure, personal appeals and coalition tactics. The use of social tactics to influence one another is affected by status differences. It is less plausible to imagine low-status members pressuring high-status members by making demands or threatening them, than vice versa. They will probably try to persuade high-status people by using factual information or flattery (Levi, 2007; Yukl, 1989).

On a more unconscious level, group negotiation is liable to the use of *cognitive heuristics*. On the one hand these heuristics accelerate efficient decision-making, on the other hand they can undermine the quality of its outcomes. Kahneman and Traversky (1973) identified three types of cognitive heuristics which enable people to understand their complex environment, availability, representativeness and anchoring. *Availability* relates to the inclination to rely predominantly on information that is most salient in one's memory. The *representativeness* heuristic refers to the tendency to judge something or someone based on its most evident features. *Anchoring* involves the strong reliance on randomly determined anchoring points. Often opening statements serve as point of reference for all statements being made thereafter. With regard to panel review this implies the comments of the first reviewer are very influential and set the tone for further discussion (Van Arensbergen et al., *forthcoming*). Knowledge of these cognitive heuristics can be implemented as social tactics when panelists actively use them to influence negotiation outcomes.

So far we have described how outcomes of panel reviews are based on characteristics of the object (features of the applicant and application) and how panel review processes are affected by social characteristics of this type of group interaction (panel composition, group norms and cohesiveness, information distribution and motivation). How these latter factors influence review outcomes, is a major task of research within the field of SoS/SPS. Finally we will focus on the context of review processes. How do review procedures and external contextual factors influence review outcomes?

## 3.6 The organizational context of panel review

The organization of panel reviews involves various aspects. For example, selecting and installing panelists, determining specific panel tasks, developing review procedures and guidelines, and implementing some sort of control mechanism, e.g. by the presence of independent supervisors or by having to write detailed review reports. Then there are contextual factors related to review processes like available budget and time pressure. With regard to panel composition we showed earlier how it may affect evaluation processes. For example, composition is found to determine the representation of expertise, social identities, status and interests among its members. These constellations can impact communication, behavior and information sharing within groups, subsequently affecting panel results.

The general task of review panels as discussed in this paper is to evaluate the quality of scientific work, of research proposals, or of scientists. The actual objects of review processes can therefore vary from hard copy research proposals and curriculum vitae to people in one's own proper person. Review procedures are designed accordingly and panels are generally clearly instructed how to execute their task. The presence of clear decision-making procedures decreases the risk of groupthink (Esser, 1998). However, procedural rules and *guidelines* generally do not fully steer review processes. Behavior of panelists can for example be more susceptible for (implicit and explicit) group norms (Spector, 1996). Langfeldt (2001), who observed panel meetings of the Norwegian Research Council, furthermore found that although review protocols prescribed quality criteria to be used, the weight assigned to these criteria differed within and between panels. The kind of criteria eventually used by panelists depended strongly on *budget* restrictions and *rating scales* they had to use. For example tight budgets and fine-rating scales tend to strengthen established research and allow less pluralism in funded research. At the same time she

found that reviewers who individually reviewed applications and send their reviews per mail, more consciously attempted to follow guidelines than panelists who discussed applications in panel meetings. This underlines the limited effect of guidelines in processes of social interaction.

With regard to funding decisions, often panels have to judge all applications, resulting in rank orders and subsequently in selections of the 'best'. This final selection decision can be made by panel themselves or by organizing parties like research councils. Consulting (in writing) external experts for every specific application can also be part of the procedure. The specific task assigned to panels is found to affect decision-making processes (Hollingshead, 1996; Langfeldt, 2002; Stewart & Stasser, 1998). In an experiment Hollingshead (1996) studied the effect of group decision-making procedures on information sharing. She compared groups who were instructed to rank all alternatives to those who had to choose the best alternatives. Ranking groups were found to exchange more information and to consider all alternatives, eventually taking better decisions than selecting groups. However, the beliefs panelist have about the correctness of their decisions also play a role. When people believe there is only one correct answer (solve task) instead of no correct answer (judge task), they tend to produce more discussion (Stewart & Stasser, 1998). According to Langfeldt (2002) also explicitness of review procedures influences decision-making processes. She distinguishes between sounding and open confrontation as two ends of a continuum. Sounding involves tacit exploration of opinions, no explicit voting and an emphasis on reaching consensus, whereas explicit voting without any preceding exploration of opinions is called open confrontation. The open confrontation method may be more efficient in terms of time needed for decision-making, but at the same time may have rather negative consequences for group cohesiveness. With regard to explicit voting, it is good to realize that the timing and sequence of voting calls influences the preferences of panelists (Davis et al., 1988).

Another aspect of review procedures is the degree of accountability of panels regarding their decisions. Accountability increases epistemic motivation of individuals, their need to search for information, the extent to which they repeat unshared information, and the quality of decision-making (De Dreu et al., 2007; Scholten et al., 2007). Therefore it is an important factor influencing chances on groupthink (Esser, 1998). Groupthink is more likely to occur in groups where any degree of accountability is absent. Making individuals accountable is found to be more effective on reducing groupthink tendencies than making them accountable collectively as a panel (Kroon, 't Hart, & Van Kreveld, 1991).

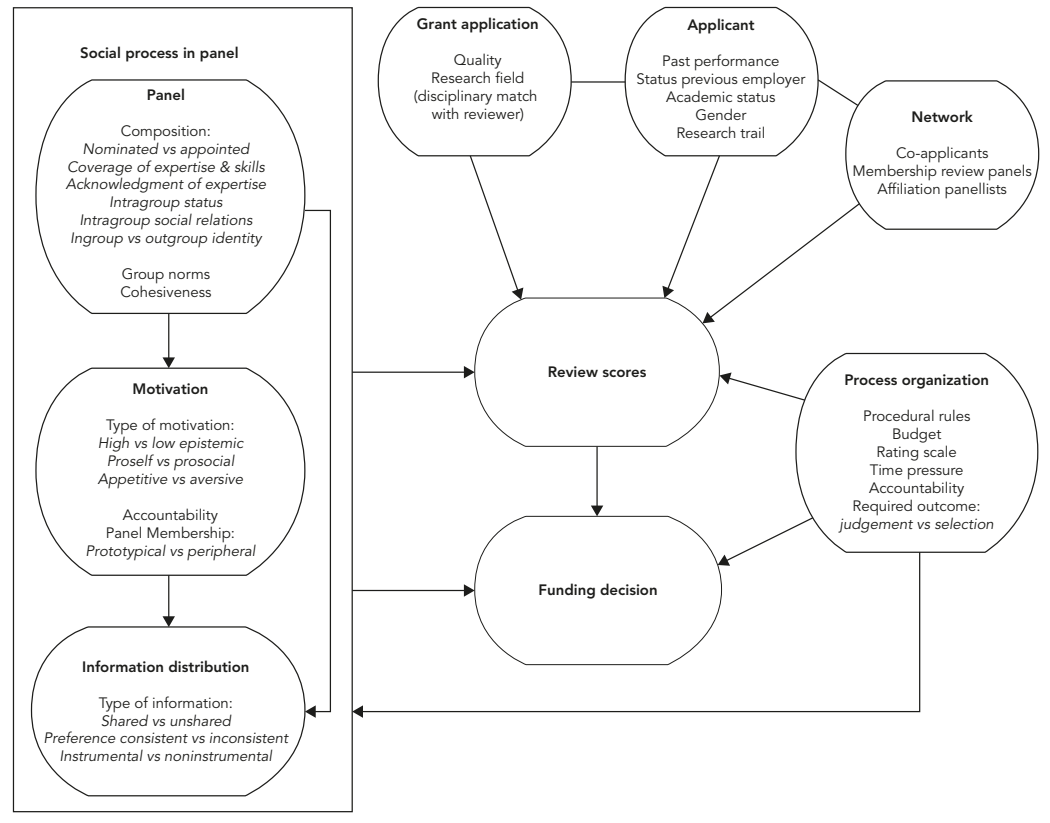
A last important factor usually seen as complicating review processes is the available time. A thorough review of all applications generally requires a lot of time, which is often at the expense of valuable research time of panelists. The combination of the large scope of applications to be evaluated and the restricted time available, reduces the ambitions of panelists to execute very rigorous reviews (De Dreu et al., 2007; Langfeldt, 2002). When panels experience strong time pressure reviewers pay more attention to shared information and less to alternatives, consequently resulting in a *closing of the mind*. People tend to rely more on cognitive heuristics like the availability heuristic as mentioned earlier and are more focused on reaching (cognitive) closure (De Dreu et al., 2007). Therefore, high time pressure is considered an important antecedent for groupthink (Janis, 1982).

## 3.7 Conclusions and implications

This paper aims to contribute to a better understanding of panel review processes by combining literature from the fields of SoS and SPS with SP. Considering the complexity of review processes characterized by social interaction, innumerable factors can be identified influencing review processes and their outcomes. Since it is impossible to include them all, we chose to focus on a limited number of factors we consider to be most illustrating based on several observations of panel meetings.

Figure 2 depicts a more detailed model of grant allocation processes, including the main social psychological factors discussed. Some of these factors - mainly related to applicants and proposals - are found to influence review outcomes too. Other factors - mainly related to panels and social interaction - need further research to determine their effect on outcomes of review processes. These factors (within the rectangle) contribute to the uncertainty with which review outcomes can be predicted outright using criteria related to scholarly quality. As shown, there are many non (directly) quality related criteria involved in review processes.

**Figure 2** More detailed model of grant allocation processes



Despite the need for further research, we may be able to formulate recommendations on how to stimulate open and thorough panel discussions resulting in fair and good quality outcomes - based on general SP studies reviewed above.

As groupthink leads to flaws in decision-making, situations facilitating groupthink should be avoided. Therefore panels should be heterogeneous with regard to their composition. At the same time an overlap in competences is recommended, as this is associated with better cooperation between panelists and allows for open discussion between experts. However, in order to benefit from diverse panel compositions, this diversity should be evident to panelists. They should be aware of the specific expertise of individual members, so they can value their information accordingly. Cohesiveness was found to be beneficial for team spirit and good communication. Therefore it is important to maintain a sense of collegiality and a good atmosphere during panel meetings. At the same time, there should be room for deviating opinions and open conflict between reviewers in order to stimulate information sharing. Next, time pressure should be reduced as much as possible, as this negatively impacts review processes. Of course this is easier said than done, however it is important to provide reviewers with sufficient time and resources to successfully do their job. Making panelists more accountable will reduce the risk of groupthink, as it has a positive effect on searching for and sharing of information and it prevents panels from wanting to reach decisions too easily and prematurely. Having to report on review processes, also involves being more explicit on which criteria were used. On the one hand this might decrease the impact of self-interests, and on the other hand it might increase the extent to which applications are reviewed in a similar way. For example by using detailed protocols including explicit standards for academic competence, decisions can be expected to be less influenced by double standards, non-instrumental information or personal interests. Another method is to assign an independent chair or moderator to each panel. This person should not be involved in review processes concerning the content, but should watch over review processes and should explicitly pay attention to types of information being discussed and weight given to it. So for example, point out the irrelevance of non-instrumental information, prevent that unique or preference inconsistent information is neglected, or too much weight is assigned to relatively unimportant information known by everyone. In order to avoid reaching agreement too quickly without considering relevant alternatives, (s)he could take on the role of devil's advocate. This will involve reviewers in alternative ways of thinking and forces them to explicitly justify their decisions. The person guarding review processes should also make sure all panelists get the chance to express their opinions, disregarding perceived differences in status or social identity - as these factors cannot be easy moderated in another way.

## 3.8 Future Research

Future research is necessary to test how factors related to processes of group decision-making as described above, affect outcomes of panel reviews. More specifically, the role of panel composition, motivation of panelists, type of information being distributed and exchanged, and accountability need further investigation. As does the role of implicit criteria deployed by panelists - compared to formal criteria as specified in official selection procedures. This future research on panel reviews should therefore also deploy observational research methods. Direct observations are hardly ever carried out. When the aim is to really understand review processes this can be considered a methodological problem, since results are now mainly based on indirect reconstruc-

tions of review processes. Ethnographic observation is therefore desirable as it offers opportunities to investigate review and decision making processes where they happen. We plan to do observation studies in a just started project.[8]

## 3.9 References

Baron, R. S. (2005). So right it's wrong: Groupthink and the ubiquitous nature of polarized group decision-making. *Advances in Experimental Social Psychology, 37*, 219-253.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology, 75*, 19-32.

Bastardi, A., & Shafir, E. (2000). Nonconsequential reasoning and its consequences. *Current directions in psychological science*, 9(6), 216-219.

Bazeley, P. (1998). Peer review and panel decisions in the assessment of Australian Research Council project grant applicant: what counts in a highly competitive context? *Higher Education, 35,* 435-452.

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003).Cohesion and performance in groups: a meta-analytic clarification of construct relations. *Journal of Applied Psychology, 88*(6), 989-1004.

Bonner, B. L., Baumann, M. R., & Dalal, R. S. (2002). The effects of member expertise on group decision-making and performance. *Organizational Behavior and Human Decision Processes, 88*, 719-736.

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology 45*(1), 197-245.

Bornmann, L., & Daniel, H.-D. (2005). Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees' decisions. *Scientometrics, 63*(2), 297-320.

Bornmann, L., & Daniel, H.-D. (2006). Selecting scientific excellence through committee peer review. *Scientometrics, 68*(3), 427-440.

Bornmann, L., Leydesdorff, L., & Van den Besselaar, P. (2010). A Meta-evaluation of Scientific Research Proposals: Different Ways of Comparing Rejected to Awarded Applications. *Journal of Informetrics 4*, 211-220.

Bornmann L., Mutz R., and Daniel H.-D. (2007). 'Gender differences in grant peer review: A meta-analysis'. *Journal of Informetrics, 1*, 226-238.

Bornmann, L., Wallon, G., & Ledin, A. (2008). Does the committee peer review select the best applicants for funding? An investigation of the selection process for two European Molecular Biology Organization programmes. *PLoS One, 3*(10), e3480.

Brouns, M. (2000). The gendered nature of assessment procedures in scientific research funding:The Dutch case. *Higher Education in Europe, 25*(2), 193-199.

Cañibano, C., Otamendi, J., & Andújar, I. (2009). An assessment of selection processes among candidates for public research grants: the case of the Ramón y Cajal Programme in Spain. *Research Evaluation, 18*(2), 153-161.

Ceci, S.J. & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *PNAS, 108*(8), 3157-3162.

8 Within the GendERC project.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science, 214*, 881-886.

Davis, J. H., Stasson, M., Ono, K., & Zimmerman, S. (1988). Effects of straw polls on group decision making: Sequential voting pattern, timing, and local majorities. *Journal of Personality and Social Psychology, 55*(6), 918-926.

De Dreu, C. K. W., Beersma, B., Steinel, W., & Van Kleef, G. A. (2007). The psychology of negotiation: Principles and basis processes. In A. W. Kruglanski & E. T. Higgins (Eds.), *Handbook of basic principles in social psychology* (2nd ed., pp. 608-629). New York: Guilford.

De Dreu, C. K. W., Beersma, B., Stroebe, K., & Euwema, M. C. (2006). Motivated information processing, strategic choice, and the quality of negotiated agreement. *Journal of Personality and Social Psychology, 90*(6), 927-943.

De Dreu, C. K. W., Nijstad, B. A., & Van Knippenberg, D. (2008). Motivated information processing in group judgement and decision making. *Personality and Social Psychology Review, 12*, 22-49.

De Jonge Akademie (2010). *Rendement van talent. Aanbevelingen voor motiverend en stimulerend loopbaanbeleid*. Amsterdam: De Jonge Akademie.

Esser, J. K. (1998). Alive and well after 25 years: A review of groupthink research. *Organizational Behavior and Human Decision Processes, 73*(2/3), 116-141.

Foschi, M. (2004). Blocking the use of gender-based double standards for competence. In M. Brouns & E. Addis (eds.). *Gender and excellence in the making* (EUR 21222 Report). Brussels: European Commission.

Gigone, D., & Hastie, R. (1993). The common knowledge effect: Information sharing and group judgment. *Journal of Personality and Social Psychology, 65*(5), 959-974.

Hartmann, I., & Neidhardt, F. (1990). Peer review at the Deutsche Forschungsgemeinschaft. *Scientometrics, 19*(5-6), 419-425.

Hodgson, C. (1997). How reliable is peer review? An examination of operating grant proposals simultaneously submitted to two similar peer review systems. *Journal of Clinical Epidemiology, 50*, 1189-1195.

Hogg, M. A., & Terry, D. J. (2000). Social identity and self-categorization processes in organizational contexts. *The Academy of Management Review, 25*(1), 121-140.

Hollingshead, A. B. (1996). The rank-order effect in group decision making. *Organizational Behavior and Human Decision Processes, 68*(3), 181-193.

Hornborstel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & von Ins, M. (2009). Funding of young scientists and scientific excellence. *Scientometrics, 79*(1), 171-190.

Janis, I. L. (1982). *Groupthink. Psychological studies of policy decisions and fiascoes* (second ed.). Boston: Houghton Mifflin Company.

Jayasinghe, U. W., Marsh, H. W., & Bond, N. (2001). Peer review in the funding of research in higher education: The Australian experience. *Educational Evaluation and Policy Analysis, 23*(4), 343-364.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237-251.

Kramer, R. M., Pommerenke, P., & Newton, E. (1993). The social context of negotiation: Effects of social identity and interpersonal accountability on negotiator decision making. *Journal of Conflict Resolution, 37*(4), 633-654.

Kroon, M.B.R., 't Hart, P., & Van Kreveld, D. (1991). Managing group decision making processes: Individual versus collective accountability and groupthink. *International Journal of Conflict Management, 2*(2), 91-115.

Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment*. Cambridge/London: Harvard University Press.

Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science, 31*(6), 820-841.

Langfeldt, L. (2002). *Decision-making in expert panels evaluating research. Constraints, processes and bias*. Norwegian Institute for Studies in Innovation, Research and Education, Oslo.

Laudel, G. (2006). The 'quality myth': Promoting and hindering conditions for acquiring research funds. *Higher Education, 52*, 375-403.

Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C.R. (2013). Global gender disparities in science. *Nature, 504*, 211-213.

Levi, D. (2007). *Group dynamics for teams* (Second ed.). Los Angeles: Sage Publications.

Marques, J. M., Abrams, D., Paez, D., & Hogg, M. A. (2001). Social categorization, social identification, and rejection of deviant group members. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (Vol. 3, pp. 400-424). Oxford, UK: Blackwell Publishers.

Marsh, H.W., Bornmann, L., Mutz, R., Daniel, H.-D., & O'Mara, A. (2009). Gender Effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research, 79*(3), 1290-1326.

Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications. Reliability, validity, bias, and generalizability. *American Psychologist, 63*(3), 160-168.

Melin, G., & Danell, R. (2006). The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy, 33*(10), 702-712.

Merton, Robert K., 1973 [1942]. The Normative Structure of Science. Robert K. Merton (ed.), *The Sociology of Science*. Chicago: The University of Chicago Press, 267-278.

Mojzisch, A., Grouneva, L., & Schulz-Hardt, S. (2010). Biased evaluation of information during discussion: Disentangling the effects of preference consistency, social validation, and ownership of information. *European Journal of Social Psychology, 40*, 946-956.

Mojzisch, A., Schulz-Hardt, S., Kerschreiter, R., Brodbeck, F. C., & Frey, D. (2008). Social validation in group decision-making: Differential effects on the decisional impact of preference-consistent and preference-inconsistent information. *Journal of Experimental Social Psychology, 44*, 1477-1490.

Mutz, R, Bornmann, L. & Daniel, H.-D. (2012). Does gender matter in grant peer review? An empirical investigation using the example of the Austrian Science Fund. *Zeitschrift für Psychologie, 220*(2), 121-129.

NWO (2012). NWO *Jaarverslag 2012*. Den Haag: NWO.

Olbrecht, M. & Bornmann, L. (2010). 'Panel peer review of grant applications: what do we know from research in social psychology on judgment and decision-making in groups?'. *Research Evaluation, 19*(4): 293-304.

Peterson, R. S., Owens, P. D., Tetlock, P. E., Fan, E. T., & Martorana, P. (1998). Group dynamics in top management teams: Groupthink, vigilance, and alternative models of organizational failure and success. *Organizational Behavior and Human Decision Processes, 7*3(2/3), 272-305.

Porter, A. L., & Rossini, F. A. (1985). Peer review of interdisciplinary research proposals. *Science, Technology & Human Values, 10*(3), 33-38.

Ranga, M., Gupta, N. & Etzkowitz, H. (2012). *Gender effects in research funding: a review of the scientific discussion on the gender-specific aspects of the evaluation of funding proposals and the awarding of funding*. Bonn: Deutsche Forschungsgemeinschaft.

Rinia, E.J., Van Leeuwen, Th.N., Van Vuren, H.G., & Van Raan, A.F.J. (2001). Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research. *Research Policy, 30*, 357-361.

Sändstrom, U., & Hallsten, M. (2008). Persistent nepotism in peer-review. *Scientometrics, 74*(2), 175-189.

Scholten, L., Van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. W. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology, 43*, 539-552.

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology, 91*(6), 1080-1093.

Spector, P. E. (1996). *Industrial and organizational psychology. Research and practice*. New York: John Wiley & Sons, Inc.

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology, 31*, 244-265.

Stephan, P. (2012). How economics shapes science. Harvard University Press: Cambridge.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly, 63*(3), 224-237.

Stewart, D. D., & Stasser, G. (1998). The sampling of critical, unshared information in decision-making groups: The role of an informed minority. *European Journal of Social Psychology, 28*, 95-113.

Ten Velden, F. S., Beersma, B., & De Dreu, C. K. W. (2011). When competition breeds equality: effects of appetitive versus aversive competition in negotiation. *Journal of Experimental Social Psychology, 47*(6), 1127-1133.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behavior relationships: A role for ingroup norms. *Personality and Social Psychology Bulletin, 22*, 776-793.

Tindale, R. S., Meisenhelder, H. M., Dykema-Engblade, A. A., & Hogg, M. A. (2001). Shared cognitions in small groups. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 1-30). Oxford: Blackwell Publisers.

Travis, G. D. L., & Collins, H. M. (1991). New light on old boys: cognitive and institutional particularism in the peer review system. *Science, Technology & Human Values, 16*(3), 322-341.

Van Arensbergen, P., Hessels, L.K. & van der Meulen, B. (2013). *Talent Centraal. Ontwikkeling en selectie van wetenschappers in Nederland*. Den Haag: Rathenau Instituut.

Van Arensbergen, P., Van der Weijden, I., & Van den Besselaar, P. (*forthcoming*). Different views on scholarly talent. What are the talents we are looking for in science? *Research Evaluation*.

Van den Besselaar, P. (2012). Grant committee membership: service or self-service? *Journal of Informetrics, 6*, 580-585.

Van den Besselaar, P. (2013). More competition, better science - on the predictive validity of grant selection. In S Hinze, et. al., (eds), *Translational twists and turns: Science as a socio-economic endeavor*; Proceedings 18th STI, Berlin, 385-392

Van den Besselaar, P. & Leydesdorff, L. (2007). *Past performance as predictor of successful grant applications*. Den Haag: Rathenau Instituut.

Van den Besselaar, P. & Leydesdorff, L. (2009). Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation, 18*(4), 273-288.

Van den Brink, M. (2009). *Behind the scenes of science. Gender practices in the recruitment and selection of professors in the Netherlands*. PhD thesis, Radboud Universiteit Nijmegen.

Van der Weijden, I. & Calero Medina, C. (2014). Gender effects on evaluation indicators. Leiden: CWTS ACUMEN Deliverable. February 2014.

Van Kleef, G. A., Steinel, W., Van Knippenberg, D., Hogg, M. A., & Svensson, A. (2007). Group member prototypicality and intergroup negotiation: How one's standing in the group affects negotiation behaviour. *British Journal of Social Psychology, 46*, 129-152.

Van Leeuwen, T.N & Moed, H.F. (2012). Funding decisions, peer review, and sci*entific excellence in physical sciences, chemistry and geosciences.* Research Evaluation, 21, 189-198.

Viner, N., Powell, P., & Green, R. (2004). Institutionalized biases in the award of research grants: a preliminary analysis revisiting the principle of accumulative advantage. *Research Policy, 33*, 443-454.

Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature, 387*, 341-343.

Winquist, J. R., & Larson, J. R., Jr. (1998). Information pooling: When it impacts group decision making. *Journal of Personality and Social Psychology, 74*(2), 371-377.

Wittenbaum, G. M., Hollingshead, A. B., & Botero, I. C. (2004). From cooperative to motivated information sharing in groups: Moving beyond the hidden profile paradigm. *Communication Monographs, 71*, 286-310.

Yukl, G. (1989). Managerial leadership: A review of theory and research. *Journal of Management, 15*(2), 251-289.

Zinovyeva, N., & Bagues, M. (2010). Does gender matter for academic promotion? Evidence from a randomized natural experiment. Working papers, 2010-15, FEDEA.

# 4 Academic talent selection in grant review panels[1]


## Abstract

Career grants are an important instrument in selecting and stimulating the next generation of researchers. Earlier research has mainly focused on the relation between past performance and success. In this study, we investigate the evidence of talent and how the selection process takes place. More specifically, we investigate which quality dimensions dominate, and how changes in weighing these criteria affect the talent selection. We also study which phases in the process are dominant. Finally, we look at the effect of the gender composition of the panel on the selection outcomes. Using a dataset of the scores of 897 career grant applications, we found no clear 'boundaries of excellence' and only a few granted talents are identified as top talents based on outstanding reviews compared to the other applicants. Quite often, the scores that applicants receive change after the interview, indicating the important role of that phase. The evaluation of talent can be considered to be contextual, as the rankings of applicants changed considerably during the procedure and reviewers used the evaluation scale in a relative manner. Furthermore, talent was found to have different (low correlated) dimensions. We also found that external peer reviews barely influence the decision-making. Finally, we found no gender bias in the decisions.


## 4.1 Introduction

Attracting and maintaining well-qualified staff is essential for organizations that want to improve their status and reputation. Therefore, universities and research councils aim to select the most talented young researchers, using explicit and also often implicit criteria (Van den Besselaar & Leydesdorff, 2009). As academic career opportunities are by far outnumbered by young researchers who hope to establish an academic career (Huisman et al., 2002; Van Balen, 2010), there is strong competition among researchers (De Grande et al., 2010). Securing a personal career grant seems increasingly crucial for a successful academic career. Besides the necessary resources to conduct research, it provides recognition of one's talent by the scientific community. As both the quality of the research system and the careers of individual researchers depend on these selection processes, it is important to understand how they function.

Most research on grant allocation focuses on the outcomes, searching for predictors for success. The internal selection mechanism has barely been studied and we therefore do not know what happens during the selection process (Bornmann et al., 2010). Only a few studies have been conducted into the individual steps of the selection process (e.g., Hodgson, 1995; Bornmann et al., 2008). Bornmann et al. (2008) applied a latent Markov model to grant peer reviews of doctoral and postdoctoral fellowships. Their model shows that the first stage of the selection procedure - external review - is of great importance for the final selection decisions. External

1 This chapter has been published as Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (2014). Academic talent selection in grant review panels. In: Prpic, K., Van der Weijden, I., Aseulova, N. (eds). (Re)searching scientific careers. St. Petersburg: IHST/RAS & SSTNET/ESA.

reviews had to be positive for fellowship applicants to have a chance of being approved. However, Van den Besselaar and Leydesdorff (2009), using a different method, could not confirm this. Moreover, no correlation was found between the decision and the external review score among the top 50% of the applicants.

In this paper, we study the process of selecting scientific talent through career grants. We will show how the selection proceeds through various phases, how consistent these phases are with each other, and which phases and criteria are decisive for the final selection. We will also look at the differences between disciplinary domains and between the three grant schemes under study.

## 4.2 Theoretical background

Although 'scientific excellence' and 'talent' are commonly used (Addis & Brouns, 2004), the meaning of these concepts is contested (Hemlin, 1993). Much debated is, e.g., whether talent is innate or acquired. Talent has been explained by innate factors (e.g., Gross, 1993; Baron-Cohen, 1998), but this research is often criticized as being mainly anecdotal and retrospective (Ericsson et al., 2007). Talent is also conceived in terms of personality (and its genetic components) effecting scientific performance (e.g., Busse & Mansfield, 1984; Feist, 1998; Feist & Barron, 2003). However, others claim that people are not born to be a genius (Howe et al., 1998), as excellence is mainly determined by environmental factors, including early experiences, training, preferences and opportunities. If that is the case, talent should not be considered as a quality in itself but more as innate *potential*. Talent is a process that enhances training and, with that, performance. It involves domain-specific expertise (Simonton, 2008). Consequently, it is difficult to decide who is a talented researcher and who is not.

### 4.2.1 Peer review

Selection panel members review and discuss grant proposals or job applications and jointly identify the best ones - often using peer review reports. This decision-making process entails, among other things, reference to one's expertise, explanations of preferences, discussion between proponents and opponents, obedience (or not) to procedures and rules and, finally, reaching agreement. To study this process of scientific reviewing and decision-making, different theoretical approaches can be used. A well-known approach which prescribes how scientists should behave according to the norms and values of science - the so called 'ethos of science' - is the Mertonian sociology of science (Bornmann, 2008). One of these norms is *universalism*, which means that the judgement of knowledge claims should be based on scientific criteria only, without interference by the personal or social backgrounds of the reviewed and reviewers (Merton, 1973 [1942]). Applied to talent selection, access to scientific careers should be based on scholarly competence alone. In this context, talent relates mainly to scientific excellence.

However, Lamont (2009) describes this type of evaluation as a social, emotional and interaction process. In an observation study of grant review panels, she shows that scientific excellence does not mean the same to everyone. Panel members from different fields, with a variety of motivations, use different criteria. And even within fields, people define excellence in various ways. As excellence is not the same for everyone but rather subject to discussion and (dis)agreement, one might consider talent to be 'socially constructed' (Smith, 2001). More generally, emerging with criticism of the Mertonian sociology of science, social constructivism poses that scientific knowledge and

judgement thereof is constructed through interpretations, negotiations and accidental events (Knorr-Cetina, 1981). Cole (1992) used some elements of the constructivist approach to make a distinction between the research frontier and the core of scientific knowledge. The frontier consists of new work which is in the process of being evaluated by the community. The core involves a small number of contributions which are accepted by the community as important and true. In this respect, there is a low level of consensus on frontier knowledge and a high level of consensus on core knowledge.

Even within the Mertonian norms, grant applications (and job applicants) are not evaluated and selected separately but in comparison to competing applications (Smith, 2001). Quality is socially and contextually defined from a specific point of reference that evolves during the evaluation process (Lamont, 2009). As a result of this *contextual ranking*, one might expect that the same grant application can be valued differently across panels, process phases and time. This is exactly what Cole and Cole (1981) found in their study of the reviewing of applications for research grants from the National Science Foundation (NSF). After reviewing all and selecting half of the applications, a second group of peers reviewed and ranked the same set again. The two rankings differed substantially. Several proposals that were rejected by the NSF would have been granted if the selection had been based on the second ranking. What, then, determines whether one proposal is evaluated as being better than another? How is talent selected within peer and panel review?

Engaging peers is essential, as they are best suited to reviewing the work of 'colleagues' within their speciality (Eisenhart, 2002). However, peers are often close to the applicants, and this creates tension between peer expertise and impartiality (Eisenhart, 2002; Langfeldt & Kyvik, 2011). This relates to another tension: peer reviews ought to be neutral but not scholarly neutral. Personal interests should be eliminated and the evaluation should be based on scholarly discretion. But where are the boundaries? A third tension exists between unanimity and divergence. Grant review panels are expected to reach a unanimous decision, but at the same time divergence is considered of great value. Divergent assessments lead to discussion and contribute to the dynamics of science (Langfeldt & Kyvik, 2011). As scientific excellence is not unambiguous but defined by reviewers and panel members in their own way, grant allocation is clearly a dynamic process.

### 4.2.2 Past performance

Earlier studies on the selection of applications focused mainly on the *past performance* of the applicant.[2] Melin and Danell (2006) compared the past performance of successful and only just unsuccessful applicants to the Swedish Foundation for Strategic Research. As the mean number of publications differed only slightly between the two groups, the awarded applicants could hardly be considered to be more productive than the rejected applicants. A study of the past performance of grant applicants in the Netherlands did find an expected difference between the tracked records of awarded and rejected applicants (Van den Besselaar & Leydesdorff, 2007, 2009). However, in comparing past performance in terms of the publications and citations of the

2 For a more elaborate literature review of the process of grant reviewing and group decision-making, see: Van Arensbergen et al. (*forthcoming*), Olbrecht and Bornmann (2010). For an elaborate review of peer review including the reviewing of scientific articles, see Bornmann (2011).

awardees with the most successful rejected applicants, the latter had a slightly better average past performance than the awarded applicants. A later study found the same for German career grants (Hornbostel et al., 2009) and for international career grants in molecular biology (Bornmann et al., 2010).

In their classical study of reviews of grant applications at the NSF, Cole et al. (1981) found a weak correlation between past performance and granted funding, concluding that the allocation of grants seems to be determined about half by the characteristics of the applicant and the proposal, and about half by chance. Other research has shown that academic rank (Cole et al., 1981), research field (Laudel, 2006), the type of research (Porter & Rossini, 1985), and academic and departmental status (Cole et al., 1981; Bazeley, 1998; Jayasinghe et al., 2003; Viner et al., 2004) (weakly) correlate with quality assessments of applications or applicants. Interestingly, there is barely any literature on the predictive validity or performativity of peer review: do the selected applicants have a better *ex post* performance than the non-selected (Bornmann, 2011; Van den Besselaar, 2013), and is this because the better candidates were selected or because getting the grant produces the best researchers as they have better resources than others?

The chance element reported by Cole et al. (1981) can be partly ascribed to the subjective character of the reviewing process and the social construction of scientific quality. According to Lamont (2009), it is impossible to completely eliminate this subjectivity, given the nature of the processes. The outcomes of the review process, therefore, are affected by who is conducting the review and how the panel is composed (Langfeldt, 2001; Eisenhart, 2002; Langfeldt & Kyvik, 2011). Different mechanisms can be discerned. Firstly, panel members who are nominated by the applicants give higher ratings (Marsh et al., 2008). Secondly, relations between reviewers and applicants influence the ratings. Researchers affiliated with reviewers received better reviews than those without this type of affiliation (Sandstrom & Hallsten, 2008). Thirdly, the way in which the review process is organized influences the outcomes (Langfeldt, 2001). Finally, the importance of the gender dimension is often debated. Given the low number of females in top academic positions, and consequently the lack of female reviewers (Wennerås & Wold, 1997), as well as the persistence of the so-called 'glass ceiling', an empirical analysis is hard to come by. The available empirical evidence provides contradictory results. Broder (1993) examines the rating of proposals from the NSF and finds that female reviewers rate female-authored NSF proposals lower than do their male colleagues. A study by Zinovyeva and Bagues (2011) showed that the gender composition of committees in Spanish universities strongly affects the chances of success of candidates applying to full professorship positions but that it has no effect on promotions to associate professorships. De Paola and Scoppa (2011) did a similar study in an Italian university and showed that gender in the composition of evaluation committees does matter. In competitions in which the evaluators are exclusively male, female candidates are less likely to be promoted. However, gender discrimination almost disappears when the candidates are judged in a panel of mixed gender.

## 4.3 Data, research questions and methods

### 4.3.1 The case

Our dataset consists of 1,539 career grant applications. These involve personal grants for researchers in three different phases of their careers:

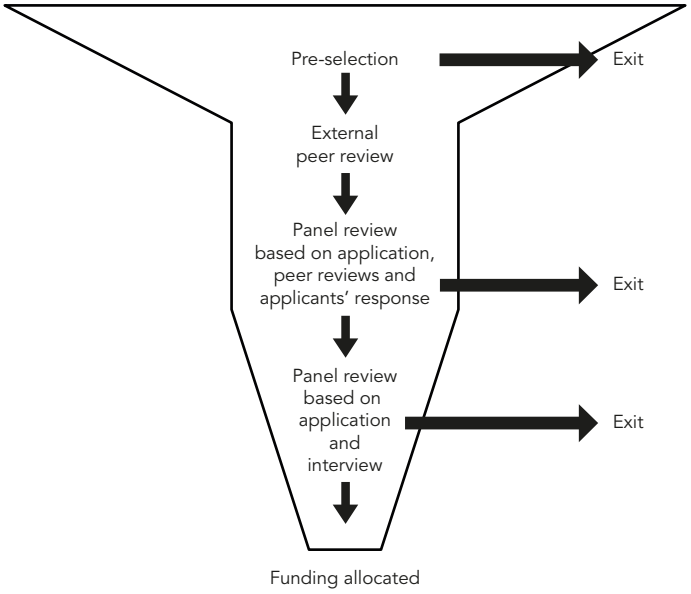
- The early career grant scheme (ECG) for researchers who have received a PhD within the previous three years. The grant offers them the opportunity to develop their ideas further.

- The intermediate career grant scheme (ICG) for researchers who have completed their doctorates with a maximum of eight years and who have already spent some years conducting postdoctoral research. The grant allows them to develop their own innovative research line and to appoint one or more researchers to assist them.

- The advanced career grants scheme (ACG) for senior researchers with up to 15 years postdoctoral experience, and who have demonstrated the ability to successfully develop their own innovative lines of research and to act as coaches for young researchers. The grant allows them to build their own research group.

Figure 1 briefly describes the selection procedure. If the number of applications in the ECG and ICG programmes is more than four times as high as the number of applications that can be awarded (as is generally the case), a pre-selection will take place - this resulted, in our case, in an overall rejection rate of about 40% of the applications, but with substantial differences between the fields. Because our dataset contains no further information on the criteria and assessments involved in the pre-selection, we do not include this phase in our study. In the ACG programme, researchers first submit a pre-proposal. The selected applicants are invited to submit a full, more detailed proposal. In addition, the selection of pre-proposals is left out of our study, for the same reasons. This reduces the dataset to 897 applications.

**Figure 1** The general grant allocation procedure.



Next, the applications are sent to external referees who are considered to be experts in relation to the research of the applicant. The number of referees varies between two and six per proposal. The reviews and the applicants' rebuttal are sent to the review panel. Partly based on this input, the panel evaluates every proposal on three criteria: the quality of the researcher (QR), the quality of the proposal[3] (QP) and research impact (RI).[4] The score on research impact is only taken into account if it is better than the proposal score.[5] When this is the case (QP<RI), the final panel score is calculated as follows:

$$\text{Total panel score} = \tfrac{1}{2}\,QR + \tfrac{1}{4}\,QP + \tfrac{1}{4}\,RI$$

If the research impact is scored lower than the quality of the proposal (i.e., if QP>RI), the panel score is calculated as:

$$\text{Total panel score} = \tfrac{1}{2}\,QR + \tfrac{1}{2}\,QP$$

The total panel score leads to a ranking of the applications which determines who proceeds to the next round: the interview, where the applicants present their proposal for the panel. Hereafter, the panel again evaluates every interviewed applicant (N = 552) on the same three criteria, taking into account the information from the previous phases. To arrive at the final panel score, the same calculation rule is used as was the case prior to the interview. The ranking of the final panel scores determines which applications will receive funding and which are rejected.

The research council consists of eight scientific divisions,[6] which are aggregated into three domains:[7] 1) the social sciences and humanities (SSH), 2) science, technology and engineering (STE), and 3) life and medical sciences (LMS). In our analyses, we will distinguish between these domains when relevant. Table 1 gives an overview of the number of applications per programme and domain. As mentioned earlier, we do not include the applications rejected in the pre-selection phase.

Our data include several attributes of the applications and applicants: gender, the grant scheme, the scientific division and the domain of the application, the referee scores, the panel scores on the three criteria, and the decisions. Between a third (ACG) and a quarter (ICG) of the applications that made it through the pre-selection received funding (table 1).

---

3 More precisely, this is the quality, innovative nature and academic impact of the proposed research.

4 This is the intended societal, technological, economic, cultural or policy-related use of the knowledge to be developed over a period of 5-10 years.

5 From 2012, the Research Impact score will always be included in the calculation of the total panel score.

6 These are the following divisions: (1) earth and life sciences (ELS); (2) chemistry (CH); (3) mathematics, computer science and astronomy (MCA); (4) physics (PH); (5) technical sciences (TS); (6) medical sciences (MS); (7) social sciences (SS); (8) humanities (HU). Around 7% of the applications are cross-divisional (CD).

7 We aggregated the scientific dimensions to the domain level as follows: SSH: social sciences and humanities; STE: chemistry, mathematics, computer sciences and astronomy, physics, and technical sciences; LMS: earth and life sciences, and medical sciences.

**Table 1** Number of applications per scientific domain and funding programme across the selection procedure.

| | | ECG | | | ICG | | | ACG | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1st review* | 2nd review# | Granted | 1st review | 2nd review | Granted | 1st review | 2nd review | Granted |
| SSH | N | 141 | 129 | 54 | 111 | 70 | 28 | 22 | 22 | 9 |
| | % | | 91.5 | 38.3** | | 63.1 | 25.2 | | 100.0 | 40.9 |
| STE | N | 151 | 70 | 40 | 124 | 65 | 33 | 34 | 34 | 12 |
| | % | | 46.4 | 26,5 | | 52.4 | 26.6 | | 100.0 | 35.3 |
| LMS | N | 161 | 76 | 49 | 118 | 56 | 28 | 35 | 30 | 10 |
| | % | | 47.2 | 30.4 | | 47.5 | 23.7 | | 85.7 | 28.6 |
| Total | N | 453 | 275 | 143 | 353 | 191 | 89 | 91 | 86 | 31 |
| | % | | 60.7 | 31.6 | | 54.1 | 25.2 | | 94.5 | 34.1 |

Rathenau Instituut

* : external reviewers & 1st panel review; # 2nd panel review
** : If we include all applications, also those rejected in the pre-selection phase, the SSH success rate is lower than the two others. This is due to the very high rejection rate in the SSH pre-selection.

### 4.3.2 Research questions

The grant allocation procedure (Figure 1) resembles a pipeline model. At the start, there is a big pool of applicants but as the procedure progresses the number of applicants decreases, with only a minority successfully reaching the end: receiving funding. In this study, we aim to understand how applications pass the selection procedure and what determines which applications are eventually successful and which are expelled along the way. This should show how talents are identified or created by the selection process. We answer the following research questions:

1. *How evident is talent?*
   How strong are the correlations between the various reviewers' scores? The stronger they correlate, the more 'evident' talent is. Secondly, do scores vary strongly? Do the selected applicants have significantly higher scores than the non-selected? Thirdly, can we clearly distinguish top talents from the other talents?

2. *Is talent selection dependent on the procedure?*
   Do the rankings of applications in the different phases of the procedure correlate? Is the result stable or does additional information in later phases result in strong fluctuations? And, are reviewers using the evaluation scales consistently throughout the procedure - do scores have a stable meaning?

3. *Which dimensions of talent can be distinguished?*
   Do the three main criteria used by the panels represent different dimensions - or do they in fact measure the same? If they are different, are the rankings dependent on weighting the dimensions? And, what does a change in weighting mean for the selection outcomes?

4. *Which phases of the process and which criteria eventually determine which applicants are considered to be talents?*

A logistic regression analysis is used to identify which criteria and phases of the selection procedure have most influence on the final grant allocation decision.

5. *Is talent gender-sensitive?*
   Does the gender composition of the panel influence the selection outcomes?

After answering these questions, we will discuss the implications of the findings for the system of selecting and granting research proposals.

### 4.3.3 Methods

Some of the following analyses are conducted on the domain and programme levels, others on the complete dataset. In the latter case, the data is standardized beforehand on the domain and programme variables. This was done by calculating the z-scores at the levels of programmes and fields.

Agreement between reviewers is analysed by calculating the standard deviation and the maximum difference between review scores per application and by rank order correlation. We will rank the review scores for each step of the selection process and compare these rankings to see whether applicants were evaluated differently at various point in the procedure. The use of the evaluation scale is analysed with chi-square tests. Rank order correlations are calculated between the three evaluation criteria used by the panels. This will show whether talent has just one or else various dimensions. Finally, to identify the predictors for talent selection we conducted a multiple logistic regression analysis.

## 4.4 Results

Evaluation practices differ between scientific domains and funding programmes (for more details, see Van Arensbergen & Van den Besselaar, 2012). Therefore, we will distinguish between the three scientific domains and funding programmes in our analyses.

### 4.4.1 The evidence of talent

The applications are refereed by external reviewers and (twice) by a panel. The number of external reviewers per proposal varies between two and six.[8] In general, there are two reviewers for the ECG, three for the ICG and four for the ACG. The external reviewers assign scores on a six point scale ranging from excellent (1), very good (2) and very good / good (3), to good (4), fair (5) and poor (6); this scale (or a similar one) is used throughout the whole procedure. We calculated the difference between the maximum and minimum review scores per proposal. As Table 2 shows, the reviewers disagree least in the ECG scheme ($M$ = 1.59; $SD$ = 1.27) and most in the ACG scheme ($M$ = 2.22; $SD$ = 1.33). The level of disagreement differs significantly between the schemes, $F(2.895) = 18.72$, $p < .001$, indicating that the further an applicant is in his career, the stronger the average disagreement about his quality.

8 Note that the applications are sent to different external reviewers, so generally reviewers are involved in the evaluation of a single application.

Taking into account that the number of reviewers varies per grant scheme, we compare the average distribution of review scores per proposal (mean standard deviation, Table 2). The standard deviation can range from 0 (if all reviewers totally agree) to 3.54 (when reviewers totally disagree). However, no significant difference was found between the programmes. Although the maximum disagreement between reviewers increases with the career phases, the mean disagreement remains the same. The higher number of reviewers in the IGC and ACG schemes explains this: the more reviewers per proposal, the smaller the weight of reviews with extreme scores.

We repeated the analysis for each of the domains to find out whether agreement on talent differs between the domains. Only in the ICG scheme did the average disagreement (standard deviation) between reviewers significantly differ between the domains ($F(2.351) = 5.25$, $p < .01$). In the ECG and ACG schemes, no significant differences were found. Finally, for all career phases the reviewers in the SSH seem to disagree more strongly than in the other domains.

**Table 2** Disagreement in evaluations by external referees per domain and funding program.

| | Early Career Grant | | Intermediate Career Grant | | Advanced Career Grant | |
|---|---|---|---|---|---|---|
| | Maximum disagreement* | Average disagreement** | Maximum disagreement* | Average disagreement** | Maximum disagreement* | Average disagreement** |
| All | 1.57 | 1.05 | 2.06 | 1.10 | 2.22 | 1.06 |
| - SSH | 1.60 | 1.13 | 2.25 | 1.21 | 2.75 | 1.28 |
| - STE | 1.68 | 1.09 | 1.76 | 0.95 | 2.03 | 0.95 |
| - LMS | 1.45 | 0.94 | 2.18 | 1.16 | 2.08 | 1.02 |

**Rathenau Instituut**

* Mean of maximum difference between review scores per application
** Mean of standard deviation review scores per application

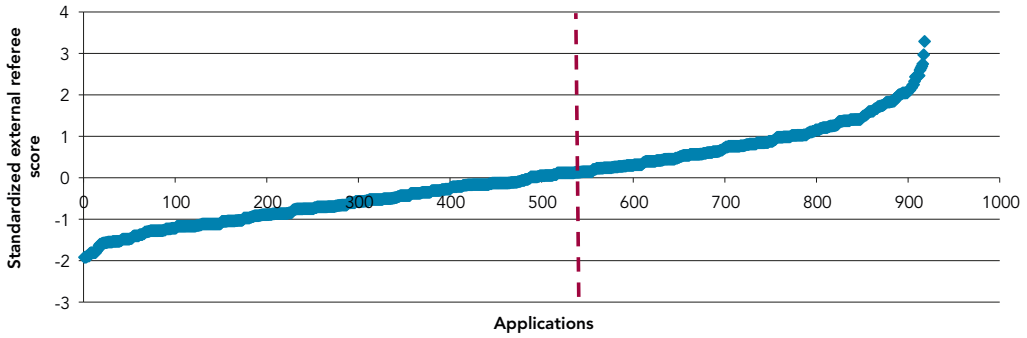
The selection of interview candidates is done by a panel, taking into account the external reviews and the applicants' rebuttal. The correlation between the standardized external review scores and the panel reviews is used to determine the extent to which evaluators in different phases of the procedure agree on the quality of applicants. In all domains, the external reviews correlate moderately strongly (ECG and ACG: $\tau = .53$, $p < .001$; ICG: $\tau = .52$, $p < .001$) with the first panel scores.[9] After the interview, the same panel evaluates the applicants again while including the new information. The correlation between the panel scores prior to and after the interview is also moderately strong in the domains of STE and LMS ($\tau = .42$, $p < .001$) and strong in SSH ($\tau = .62$, $p < .001$).

The average scores are used to distinguish between the talented and less talented applicants, but how strongly do these scores discriminate? We ranked (for the complete set and per domain) all the applications using the standardized average review score. As Figure 2 shows for the complete set, the distribution has no clear cut-off point, and a similar pattern exists at the domain and programme levels. The dotted line indicates the *de facto* cut-off point of applications

9 Since the dataset is characterized by a large number of tied ranks, we use Kendall's tau instead of Spearman's rho.

selected for the next (interview) phase. However, this selection boundary does not follow from the scores, as the difference between success and only just no success is very small. Similar patterns were found for the panel scores, where the difference between success and failure is very small too.

**Figure 2** Standardized external referee scores for the complete set of applications.



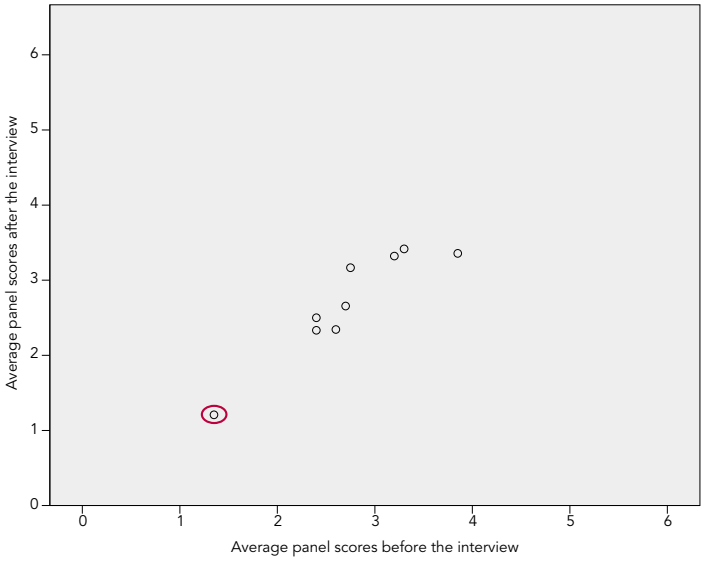

In conclusion, no clear 'boundaries of excellence' could be identified between selected and not-selected applicants. Moreover, the average scores in the three phases of the procedure only correlate moderately strongly, and this may reflect considerable changes between the rankings. This issue will be addressed in the next section, after we have looked into the evidence of top talents.

## Top talents

Figure 2 showed no clear delineation of talent but more gradual differences in talent assessment. Experienced reviewers often claim to easily identify the real top, there are always a few top talents who stand out from the beginning (Van Arensbergen, Van der Weijden & Van den Besselaar, *forthcoming*). To test this claim, we looked at the average total review scores per panel in order to identify the top talents. We determined: i) the number of positive outliers (= exceptionally high scores) in the evaluation round prior to and after the interview; ii) the distance between the outliers and the best of the gross evaluation scores; iii) the number of stable outliers (the same outliers in both evaluation rounds).

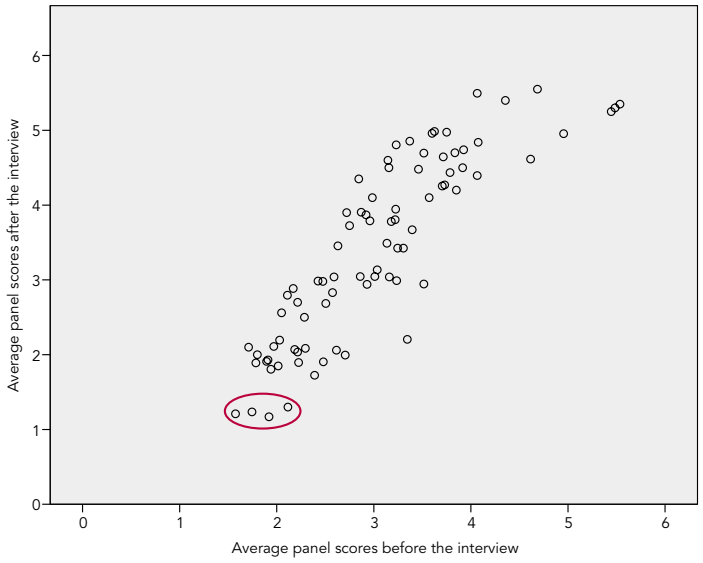
Figure 3a is an example of a panel that clearly identified a top talent both before and after the interview. Figure 3b shows that a clear top was identified only after the interview. Looking at the x-axis, the four applicants eventually identified as the top talents did not stand out in the eyes of the panel members before the interview. An example of a case where no top is recognizable but where all the applicants are close together is depicted in Figure 3c.

**Figure 3a** Average panel scores before and after interview in panel which clearly identified a top talent in both review rounds.




Rathenau Instituut

**Figure 3b** Average panel scores before and after interview in panel which identified top talents only after the interview.

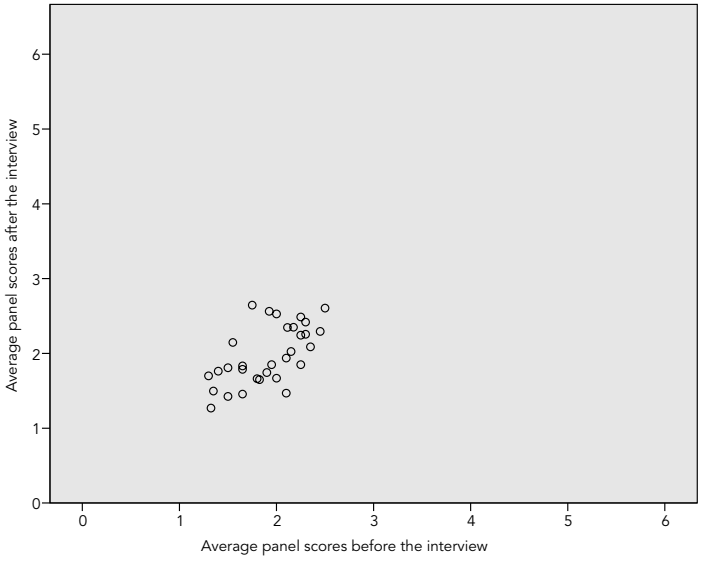



Rathenau Instituut

In general, we found that a clear top was identified more often after the interview than before (Table 3), making Figure 3b most representative for the 27 panels. In more than half of the panels, no applicants stood out from the rest before the interview, while after the interview 20 of the panels identified a top. This top predominantly consists of one person, with a maximum of four. For example, seven panels identified one top talent in the first selection phase, whereas two panels identified four top talents.

Also, after the interview, the distance between the (lowest in the) top and the (highest in the) middle group is on average a little larger (0.51, SD = 0.19) than before the interview (0.48, SD = 0.18). Panel members use an evaluation scale from 1 to 6. These average distances of 0.48 and 0.51 clearly differentiate a top from the large middle area, where there is a great deal of overlap and where most applications are very close to each other in terms of their review scores (see Figure 2).

**Figure 3c** Average panel scores before and after interview in panel which identified no top talents.




Rathenau Instituut

When we look at the stability of the top, we find that only in a few cases were the same applicants identified as top talents both before and after the interview. In 17 out of 27 panels, none of the applicants were identified as a top talent in both the evaluation rounds. In seven panels, we discerned one stable top talent. In total, of the 53 applicants who were in the top at some point of the evaluation process, 15 belonged to the top in both rounds and can be considered to be stable top talents. However, the vast majority of selected applicants (210 out of 263) were never scored as exceptional talent.

**Table 3** Number of panels (n = 27) which identified top talents before and after the interview, and which identified the same top talents in both selection phases.

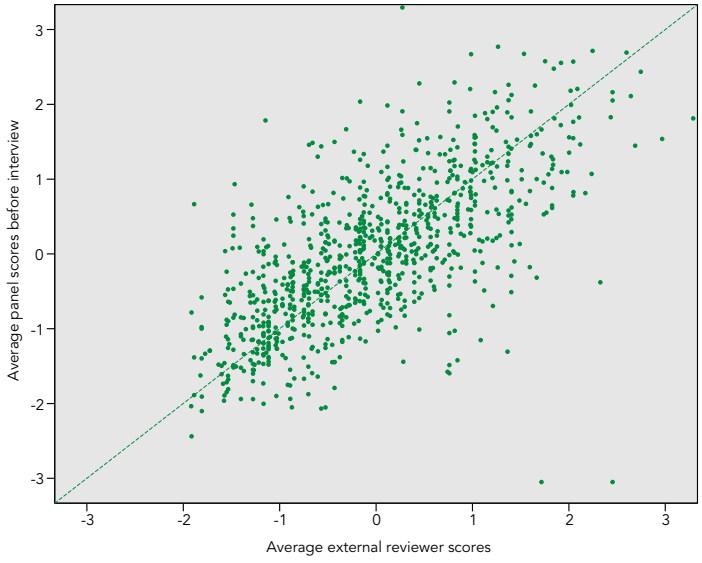
| Number of identified top talents | Before interview | After interview | Both before and after interview |
|---|---|---|---|
| 0 | 14 | 7 | 17 |
| 1 | 7 | 8 | 7 |
| 2 | 3 | 3 | 1 |
| 3 | 1 | 6 | 2 |
| 4 | 2 | 3 | 0 |

Rathenau Instituut

### 4.4.2 Effects of the procedure

The selection procedure includes three evaluation phases in which new information is added and which may influence the resulting assessment. Figures 4 and 5 show how applications are evaluated differently at different points in the procedure based on the standardized review scores.

**Figure 4** The 1st panel review by external reviewer score.



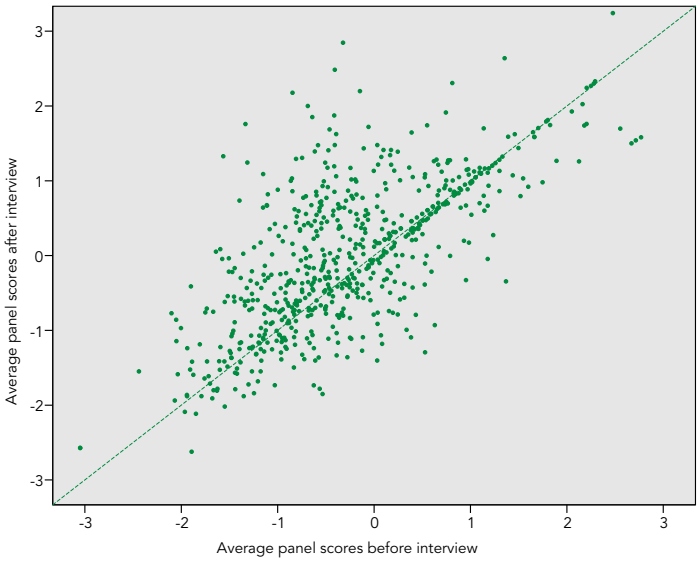

Rathenau Instituut

Right of the diagonal in Figure 4 are the applications that had a better (= lower) first panel score than external review score. On the left side are the applications that had a better external review

score. Clearly, the scores and the relative positions of the applications change during the procedure. If external (peer) review scores had been leading, the set of applicants invited to the interview would have been rather different. Since both evaluations are based on roughly the same information, this implies that talent evaluation depends upon the way in which it is organized - it is 'contextual'.

In Figure 5, the panel reviews before and after the interview are compared, with right from the diagonal those applications that scored lower (= better) after the interview than before, whereas left of the diagonal the opposite is the case. Panels adjust their assessments after the interview, and some applicants scored rather differently after the interview compared with before.

**Figure 5.** The 2nd panel review by 1st panel review.




**Rathenau Instituut**

This implies that if grant allocation had been based on the evaluation scores before the interview, the outcome would have been different. How strong is this effect? To answer this question, we compare the rankings of applications between the three evaluation moments, showing the importance of the various phases of the selection process.[10] We found that 48 (17%) of the interview candidates would not have been invited for the interview if the external referee scores had been paramount. According to the procedure, the panel score is decisive. However, there were 24 rejected applicants with a higher total panel score than the selected applicants. This means that 9% of the successful applicants were not selected simply because they were among the highest total panel scores. The panel thus in fact has additional autonomy in decision-making.

10 In some divisions and in the ACG, all the applicants were invited for the interview; these are excluded from this part of the analysis.

Grant allocation is the final step in the selection procedure. If the grant allocation had been based entirely on the evaluation by the external referees, 26% of the applicants would not have been allocated a grant. If interviews would not have been part of the procedure - and this is the case in many funding schemes - and the first panel reviews would have determined the grant allocation, 22% would have been allocated to currently unsuccessful applicants. These findings imply that the interview considerably changes the assessment of talent.[11] As the procedure prescribes, the eventual allocation decision largely corresponds to the final panel score - only 2% of the granted applicants had a lower panel score than the best rejected applicants.

Differentiating between the funding programmes and scientific domains, differences were found between domain-programme combinations but no overall pattern could be identified (for more details, see Van Arensbergen & Van den Besselaar, 2012).

## What do the scores represent?

Having shown how the *perception of talent* changed, we will now study changes in the *use of the evaluation scale* (as distinct from the evaluation of the applications). As has been said, the six point scale ranges from excellent (1), very good (2) and very good / good (3), to good (4), fair (5) and poor (6). Clearly, it is an 'absolute scale'. The panel members assign a score between one and six to each application on three criteria (quality researcher, quality proposal and research impact). Table 4 shows the mean scores and standard deviations for two typical evaluation panels both before and after the decision about which applicants are invited for an interview.

**Table 4** Use of evaluation scale.

| Case | programme & domain | | 1st panel review all applications | | | 1st panel review selected applications | | | 2nd panel review selected applications | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | researcher | proposal | total | researcher | proposal | total | researcher | proposal | total |
| 1 | ECG-STE | N | 34 | 34 | 34 | 18 | 18 | 18 | 18 | 18 | 18 |
| | | Mean | 2.89 | 3.43 | 3.10 | 2.28 | 2.87 | 2.55 | 2.56 | 3.23 | 2.84 |
| | | SD | .84 | .78 | .73 | .54 | .45 | .47 | .83 | .95 | .84 |
| 2 | ICG-SSH | N | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| | | Mean | 1.69 | 2.18 | 1.91 | 1.69 | 2.18 | 1.91 | 1.79 | 2.20 | 1.98 |
| | | SD | .38 | .43 | .35 | .38 | .43 | .35 | .38 | .49 | .39 |

Rathenau Instituut

In case one, about 50% of the highest scoring applications were selected. As expected, the means for all applicants (first review, all applications) are lower than the means for selected applicants only (first review, selected applicants).[12] The standard deviation of the entire set of applicants is larger than for the selected candidates only - indicating an expected smaller

11 In a follow-up study, we investigate the dynamics, the criteria (implicitly) applied, and the effects of the interview (Van Arensbergen et al., forthcoming).

12 Please also note that, here, lower scores correspond with higher numbers.

variation among the selected applicants. However, the average and standard deviation of the scores *after* the interview (second panel score) are equal to the values for all applicants in the first review, suggesting that the panel has again applied t*he whole scale*: some of the applications scoring very good and excellent in the first round are now only fair or even poor. In this case, the scale is used in a *relative* way, and not as an *absolute* one. In case two, no selection took place, as all the applicants were interviewed. The interview influenced individual scores, but the average and the standard deviation before and after the interview remained about the same. No changes in the use of the scale seem to have occurred here.

Comparing the 14 'selective' panels with the 12 'non-selective' panels (in Table 5) sees a significant correlation between the change of context (selection between the phases or not) and the change of the use of the scale (relative or absolute scale). Consequently, the assessment of talent depends on the context, on the procedure: e.g., an interview, as shown in the previous section, and the number of competitors, as shown in this section.

**Table 5** Changing use of the scores by changing context (n = 26).

| | | Reduction of nr applicants after 1st panel evaluation | |
|---|---|---|---|
| | | yes[a] | no |
| Decrease average score* | no | 4 (28.6%) | 10 (83.3%) |
| | yes[b] | 10 (71.4%) | 2 (16.7%) |
| Increase standard deviation** | no | 4 (28.6%) | 8 (66.7%) |
| | yes[b] | 10 (71.4%) | 4 (33.3%) |
| Total | | 14 (100%) | 12 (100%) |

Rathenau Instituut

a yes = changing context.
b yes =using the score values in a relative way.
* $X^2$=7.797, p=0.005; ** $X^2$=3.773, p=0.05.

### 4.4.3 The dimensions of talent

Earlier, we saw that the external reviews correlated moderately strongly with the panel reviews. Distinguishing between the three criteria used by the panel shows that this moderate correlation is mainly due to a relatively weak correlation between the external reviews and the panel scores for research impact, $\tau = .22$, $p < .001$ (SSH); $\tau = .29$, $p < .001$ (STE); $\tau = .36$, $p < .001$ (LMS). In the LMS domain, however, the external referee scores correlate even more weakly with the panel scores for the researcher, $\tau = .32$, $p < .001$. The external reviews are strongest in relation to the panel scores for the proposal, $\tau = .55$, $p < .001$ (SSH); $\tau = .55$ $p < .001$ (STE); $\tau = .64$, $p < .001$ (LMS).

The three criteria are found to correlate moderately with each other (Table 6). Research impact correlates most weakly with the quality of the researcher in all domains both before and after the interview, ranging from $\tau = .31$ to .41. The correlation between quality of the proposal and quality of the researcher increased after the interview in all domains, strongest in LMS, from $\tau = .44$ before the interview to $\tau = .59$ after the interview.

**Table 6** Correlations between the standardized panel review scores for the three criteria per domain.

| | | SSH | | STE | | LMS | |
|---|---|---|---|---|---|---|---|
| | | QR | QP | QR | QP | QR | QP |
| Before interview | QP | .50* | | .50* | | .44* | |
| | RI | .33* | .41* | .38* | .49* | .37* | .47* |
| After interview | QP | .57* | | .56* | | .59* | |
| | RI | .37* | .49* | .31* | .44* | .41* | .47* |

Rathenau Instituut

* QR = quality researcher; QP = quality proposal; RI = research impact
* p<.001

This suggests that the three criteria represent different dimensions. The total score of the panel (as calculated with the formulas from the method section) therefore depends on the weights attributed to the different dimensions. This may change with the decision-making context. In 2012, a change in the weighting of the research impact score was implemented in the review procedures. From now on, research impact accounts for 20% of the total panel score, and the quality of the researcher and the proposal for 40% each. We applied this new procedure to our dataset to explore how this would affect the selection outcomes.

The issue that comes up is: to what extent does the changing of weights influence the selection procedure? Would other applicants have been selected if the three criteria were weighted differently? To answer these questions, we performed some simulations in which we changed the weights. Two analyses can be done: (i) a rank order correlation between the different simulated scores informs us about the impact of the scores. The lower the rank order correlation, the greater the effect the weighting has on the resulting order of applicants. This, by the way, does not imply that changing the weight would also influence the decisions as the altered rank order may be within the set of successful applicants and within the set of unsuccessful applicants. Therefore, (ii) one should check whether the changed order would move applicants from below the success threshold to a place above the threshold, and vice versa.

## Does changing weights imply changes in the rank order?

We simulated the outcomes using five different sets of weights, as shown in Table 7. We check it here for the first decision as to whether an applicant is invited or rejected for the interview. For each of the sets, we calculated the score of the applicant, and this led to five rank orders.

**Table 7** Used weights for the three criteria.

| Weights: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Researcher | 0.5 | 0.5 | 0.33 | 0.4 | 0.4 |
| Proposal | 0.5 | 0.5 | 0.33 | 0.4 | 0.2 |
| Societal impact | 0 | + | 0.33 | 0.2 | 0.4 |

Rathenau Instituut

+: If 'societal impact' scores higher than proposal, a new value for 'proposal' is calculated as the mean of the old value of 'proposal' and the value of 'societal impact'

Using these weights, we found for the interview selection that the rank order correlations are rather high. For almost all instrument/field combinations, Spearman's Rho remained between 0.83 and 0.97 (Table 8, left part). The lowest correlations (between 0.62 and 0.80) were all between weights set at 1 (where societal impact would not be taken into account) and weights set at 5 (where societal impact would be strongly taken into account). If it is taken into account, the exact weight may not be very important for the rank order of the applications, as the correlation remains in all cases above 0.83. For the grant decision, we find a similar pattern (Table 8, right part).

**Table 8** Simulations: average correlations between rank orders based on five weights for each funding programme and field*.

| | Decisions before the interview | | | Decisions after the interview | | |
|---|---|---|---|---|---|---|
| | ECG | ICG | ACG | ECG | ICG | ACG |
| ELS | 0.93 | ** | 0.90 | 0.89 | ** | 0.97 |
| CH | 0.91 | 0.87 | 0.93 | 0.94 | 0.82 | 0.97 |
| MCA | 0.90 | 0.92 | 0.90 | 0.82 | 0.84 | 0.88 |
| CD | 0.94 | 0.90 | 0.97 | 0.95 | 0.90 | 0.83 |
| HU | ** | ** | 0.88 | ** | ** | 0.99 |
| SS | 0.84 | 0.88 | 0.84 | 0.83 | 0.92 | 0.93 |
| PH | 0.87 | 0.93 | 0.97 | 0.96 | 0.98 | 0.92 |
| TS | 0.88 | 0.88 | 0.96 | 0.89 | 0.89 | 0.99 |
| MS | 0.83 | 0.88 | 0.88 | 0.86 | 0.92 | 0.86 |

Rathenau Instituut

* We use here the more detailed division in fields (see notes 5 and 6)
** Societal impact scores not available

**What would this mean in terms of the decisions and success rates in the ECG and ICG programmes?**

Table 9 shows that the selection of grantees does depend on the selected weights. Scenario five would have changed the grant allocation between 10.4% (ECG) and 14.4% (ICG), and this is, of course, important for the involved applicants. Furthermore, the table shows that there is significant variety between the fields, as in some fields the percentage of different grantees under scenario five would be more than 50%. Independently of whether this would have had an effect on the science system, the analysis suggests that what counts as talent is indeed context dependent.

**Table 9** Scenario 5* versus scenario 2**: Number of different grantees.

| | ECG | | ICG | |
|---|---|---|---|---|
| | different grantees | % | different grantees | % |
| ELS | 1 | 5.6 | - | - |
| CH | 1 | 10.0 | 4 | 57.1 |
| MCA | 3 | 33.3 | 3 | 50.0 |
| CD | 1 | 11.1 | 2 | 28.6 |
| SS | 3 | 10.3 | 0 | 0.0 |
| PH | 0 | 0.0 | 0 | 0.0 |
| TS | 1 | 8.3 | 1 | 5.9 |
| MS | 3 | 11.1 | 1 | 5.0 |
| Total | 13 | 10.4 | 11 | 14.4 |

Rathenau Instituut

* Impact with heavy weight
** Reality (until 2012)

### 4.4.4 Predictors for talent selection

The first decision is when panels select and reject applications for the interview round, based on the external reviews, the applicants' responses to these reviews, and the panels' own scoring on three criteria. In order to determine which of these variables best predicts whether an application will be selected for the interview, we conducted a stepwise logistic regression analysis, including the average external referee score and the three panel scores.[13]

The model with only the external reviews correctly predicts in 69.1% of the cases who goes through to the interview - slightly above the random correct prediction of 61.5%. In the full model, only the panel scores for the quality of the proposal and the researcher's quality are included, whereas the other variables are excluded (Table 10). These two variables correctly predict in 77.3% of the cases whether a researcher was invited for the interview or not.

13 As the following results show, the stepwise method eliminates two variables because they do not contribute significantly to the model.

**Table 10** Logistic regression of the selection of interview candidates.

| | B (SE) | X² (df) | Nagelkerke R² | % correct |
|---|---|---|---|---|
| Constant | -0.61* (0.10) | | | |
| Quality Researcher | 0.71* (0.13) | | | |
| Quality Proposal | 1.36* (0.15) | | | |
| Model | | 283.96* (2) | .48 | 77.3 |
| Not included | | | | |
| External Reviews | 0.23 (0.16) | | | |
| Research Impact | 0.15 (0.14) | | | |

*p < .001

**Rathenau Instituut**

After the interviews, the panel again scores the applications on the three criteria. A logistic regression analysis was performed to predict the allocation decisions from the external referee scores and the three panel scores (Table 11). Again, the external referee scores and the research impact scores do not contribute significantly to the prediction. The panel scores for the proposal and for the researcher result into a correct classification in 83.1% of the cases. The model with only the external reviews correctly predicts in 65.2% of the cases who receives funding - slightly above the random correct prediction of 52.3%.

**Table 11** Logistic regression analysis to predict grant allocation decisions.

| | B (SE) | X² (df) | Nagelkerke R² | % correct |
|---|---|---|---|---|
| Constant | 0.46* (0.15) | | | |
| Quality Researcher | 1.40* (0.23) | | | |
| Quality Proposal | 1.80* (0.23) | | | |
| Model | | 294.97* (2) | .65 | 83.1 |
| Not included | | | | |
| External Reviews | 0.21 (0.18) | | | |
| Research Impact | 0.08 (0.19) | | | |

* p < .001

**Rathenau Instituut**

Distinguishing between the three funding programmes, in short we found that for early career researchers to a greater extent other factors are taken into account in decision-making. Moreover, the *de facto* weights of both of the included criteria are found to differ between the funding programmes. For the early career researchers, the evaluation of the proposal and the researcher almost evenly determine the final selection decision, whereas for the intermediate and advanced career researchers, the quality of the proposal is more important than the quality of the researcher (for more details, see Van Arensbergen & Van den Besselaar, 2012).

### 4.4.5 Is talent gendered?

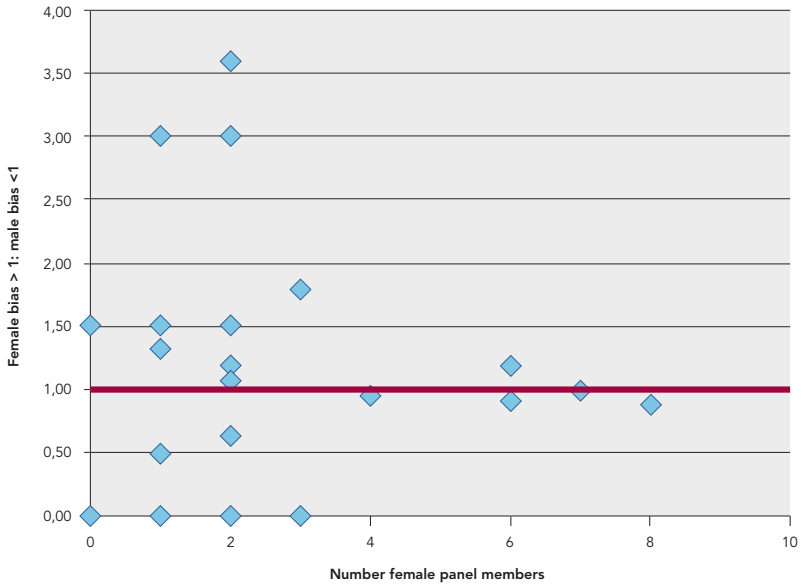
As suggested in the literature, panel composition is often found to influence decision-making: decisions of panels with no or only a few female members are found to be gender-biased. As councils increasingly claim to support female applications, it is interesting to investigate whether this effect still exists. Do 'male dominated' panels still exist and, if so, do these panels decide more often in favour of male than of female applicants? If no gender bias were to exist, then one would expect that the percentage of granted applications within the set of female applicants is similar to the percentage of granted applications in the set of male applicants. Gender bias is calculated as:

Gender bias = (success rate of women / success rate of men)

Of course, this is under the assumption that male and female applicants and applications are, on average, of equal quality. One way of tentatively testing this is by comparing the referee scores for female and male applicants. These are individually given by external reviewers - before the proposals enter the decision-making process. We found that the mean score of male applicants is slightly higher (9%) than the average score of female applicants. In most fields, this difference is not statistically significant (if we consider the data as a random sample), and insofar as the differences are significant, it is in the more advanced career schemes. For the ECG, differences are small(er) and never significant. We therefore assume that the - comparable - peer review scores are hardly gender biased - if at all (Marsh et al., 2009).

We analyse here the relation between the gender composition of panels and the final selection decision. One may do the same for the interview decision. Figure 6 shows gender bias according to the number of women in the panel. As the figure shows, there are still panels with no or only one female member. However, one cannot conclude that these panels exhibit a gender bias against female applicants. In the lower range of female panel membership, we actually find a large variation in the bias variable. If there is a pattern, it more often seems to actually be in favour of female applicants. Panels with larger numbers of female members consistently seem to have no gender bias in their decisions.

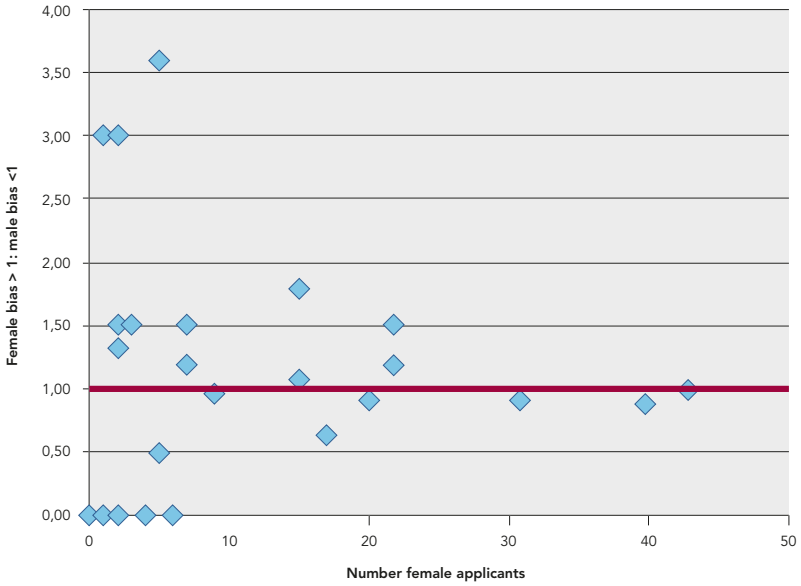
**Figure 6** Gender bias by number of female panel members.




Rathenau Instituut

Why this difference occurs needs further investigation. However, one factor may be whether a field has many or only a few female applicants. In the latter case, the success rate of women is heavily influenced by a single decision. Indeed, as Figure 7 shows, in those fields with few female applicants, the spread in the success rate is large, whereas this is not the case in those fields with many female applicants. Furthermore, one might expect that fields with only a few female applicants would also have somewhat male-dominated panels - since these fields may simply lack female researchers to occupy panels. A study of Van den Brink (2009) suggests that a gender bias in promotion decisions is due to the composition of panels. However, we cannot confirm this, as our data suggest no correlation between the number or percentage of women in a panel and the gender bias in the results.

**Figure 7** Gender bias by number of female applicants.




Rathenau Instituut

## 4.5 Conclusions and discussion

First of all, the moderate correlations between the criteria indicate that talent has different dimensions. This implies that the weight of the criteria may strongly influence the selection process. For example, the weight of research impact is very low in the case we studied, but the current tendency to include expected societal impact more strongly in the evaluation of proposals potentially leads to the selection of other types of applicants as 'the most excellent'. However, our simulations suggest that this can only happen if the weight of the societal relevance criterion were to be more substantial than currently implemented.

Secondly, the scores change considerably between the phases. Some applicants - top ranked by the external referees - are not even invited for an interview by the panels. In addition, these same panels regularly change their evaluations of applicants radically after the interview. A clear top can more often be distinguished after the interview than before; however, the actual number of identified top talents is relatively low. The interview seems decisive, but how this works needs further investigation. Does the interview bring new information, leading to a different evaluation? In that case, the procedure does influence the outcome considerably, which can of course be intended and desirable. As such, should the many existing procedures without interviews be abandoned?[14] Or is it because other aspects of talent (such as communicative skills) and several cognitive, motivational and social processes (Lamont, 2009) play a role during the interview, as well as various psychological factors (Hemlin, 2009)?

14 Interestingly, the very prestigious ERC advance grants do not include an interview with the applicants.

Thirdly, the role of the external peer review in the quality assessment seems modest (Langfeldt et al., 2010). Using only external review scores as predictors, the percentage of correctly predicted applications is only slightly higher than random (65.2% versus 52.3%) and much lower than for the two other predictors (83.1%). Combined with the moderately (but not very) high correlation ($\tau$ = .52) between review scores and panel scores, this suggests that the panel takes the review scores into account, but not very much.[15] Further study is needed, to reveal whether (and if so, how) panel members value and use the peer review reports.

Fourthly, reviewers disagree, and the further a researcher is in his/her career, the more reviewers disagree. In line with earlier studies, consensus about quality is lower in the SSH than in the natural sciences, technical sciences and life sciences (Cicchetti, 1991; Simonton, 2006). Panels and external reviewers also do not draw a clear line between talented and less-talented researchers, as for the middle group very small differences in scores eventually decide who receives a grant and who does not. As the funding decisions are of great importance for the careers of (especially) young researchers, career success becomes partly a question of luck.

Finally, the composition of the panel does not seem to result into a gender bias in the decisions. This suggests that councils' policies to stimulate female participation in science appear effective - at least at the level of their panels. Under these conditions, gender bias in outcomes seems to be related to the low number of female candidates in some fields.

In summary, our findings clearly indicate the contextuality of evaluation and decision-making. In improving the transparency, quality and legitimacy of grant allocation practices, it would therefore be desirable to analyse in more depth the details of the *de facto* (implicit and explicit) applied criteria. As the selection procedure influences the evaluation of scientific talent, we suggest using a variety of procedures instead of standardizing. The interview was found to have an important impact on the evaluation of the applicants. If communicative skills and self-confidence are decisive, the selection outcomes will be biased towards these qualities at the moment that all evaluation procedures would include interviews. Since no evident pool of talents could be identified based on the various scores, and as the differences between granted and eventually rejected applications were small, a variety of procedures may result in the selection of a variety of talent.

## 4.6 References

Addis, E. & Brouns, M. (eds) (2004). *Gender and excellence in the making*. Bruxelles: European Commission.

Baron-Cohen, S. (1998). Superiority on the embedded figures test in autism and in normal males: Evidence of an "innate talent"? *Behavioral & Brain Sciences, 21*(3), 408.

Bazeley, P. (1998). Peer review and panel decisions in the assessment of Australian research council project grant applicant: What counts in a highly competitive context? *Higher Education, 35*, 435-452.

Bornmann, L. (2008). Scientific peer review. An analysis of the peer review process from the perspective of sociology of science theories. Human Architecture: *Journal of Sociology of Self-Knowledge, 6*(2), 23-38.

15 This is in line with the findings by Hodgson (1995) and contrasts with the findings of Bornmann et al. (2008).

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science & Technology, 45*, 199-245.

Bornmann, L., Leydesdorff, L. & Van den Besselaar, P. (2010). A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. *Journal of Informetrics, 4*(3), 211-220.

Bornmann, L., Mutz, R. & Daniel, H.D. (2008). Latent Markov modeling applied to grant peer review. *Journal of Informetrics, 2*(3), 217-228.

Broder, I.E. (1993). Review of NSF economic proposals: Gender and institutional patterns. *American Economic Review, 83*(4), 964-970.

Busse, T.V. & Mansfield, R.S. (1984). Selected personality-traits and achievement in male scientists. *Journal of Psychology, 116*(1), 117-131.

Cicchetti, D.V. (1991). The reliability of peer-review for manuscript and grant submissions - a cross-disciplinary investigation. *Behavioral & Brain Sciences, 14*(1), 119-134.

Cole, S. (1992). *Making science: Between nature and society*. Cambridge, MA: Harvard University Press.

Cole, S., Cole, J.R. & Simon, G.A. (1981). Chance and consensus in peer review. *Science, 21*4, 881-886.

De Grande, H., De Boyser, K. & Van Rossem, R. (2010). Carrièrepaden van doctoraathouders in België. Loopbaanpatronen naar wetenschapsgebied. *HR2 e-document*.

De Paola, M. & Scoppa, V. (2011). Gender discrimination and evaluators gender: Evidence from the Italian academy. Italy, Department of Economics and Statistics. University of Calabria. Working paper No. 06-2011.

Eisenhart, M. (2002). The paradox of peer review: Admitting too much or allowing too little? *Research in Science Education, 32*(2), 241-255.

Ericsson, K.A., Roring, R.W. & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies, 18*(1), 3-56.

Feist, G.J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality & Social Psychology Review, 2*(4), 290-309.

Feist, G.J. & Barron, F.X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality, 37*(2), 62-88.

Gross, M.U.M. (1993). *Nurturing the talents of exceptionally gifted individuals*. In: K.A. Heller, F.J. Mönks & A.H. Passow (Eds.), *International handbook of research and development of giftedness and talent*. (pp. 473-490). Passow: Routledge.

Hemlin, S. (1993). Scientific quality in the eyes of the scientist - a questionnaire study. *Scientometrics, 27*(1), 3-18.

Hemlin, S. (2009). Peer review agreement or peer review disagreement. Which is better? *Journal of Psychology of Science & Technology, 2*(1), 5-12.

Hodgson, C. (1995). Evaluation of cardiovascular grant-in-aid applications by peer-review - influence of internal and external reviewers and committees. *Canadian Journal of Cardiology, 11*(10), 864-868.

Hornbostel, S., Bohmer, S., Klingsporn, B., Neufeld, J. & von Ins, M. (2009). Funding of young scientist and scientific excellence. *Scientometrics, 79*(1), 171-190.

Howe, M.J.A., Davidson, J.W. & Sloboda, J.A. (1998). Innate talents: Reality or myth? *Behavioral & Brain Sciences, 21*(3), 399.

Huisman, J., de Weert. E. & Bartelse, J. (2002). Academic careers from a European perspective - the declining desirability of the faculty position. *Journal of Higher Education, 73*(1), 141.

Jayasinghe, U.W., Marsh, H.W. & Bond, N. (2003). A multilevel cross-classified modelling approach to peer review of grant proposals: The effects of assessor and researcher attributes on assessor ratings. *Journal of the Royal Statistical Society Series a-Statistics in Society, 166*, 279-300.

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford, UK: Pergamon Press.

Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment.* Cambridge, MA: Harvard University Press.

Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science, 31*(6), 820-841.

Langfeldt, L. & Kyvik, S. (2011). Researchers as evaluators: Tasks, tensions and politics. *Higher Education, 62*, 199-212.

Langfeldt, L., Stensaker, B., Harvey, L., Huisman, J. & Westerheijden, D.F. (2010). The role of peer review in Norwegian quality assurance: Potential consequences for excellence and diversity. *Higher Education, 59*(4), 391-405.

Laudel, G. (2006). The 'quality myth': Promoting and hindering conditions for acquiring research funds. *Higher Education, 52*, 375-403.

Marsh, H.W., Bornmann, L., Mutz, R.D. & O'Mara, A. (2009). Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research, 79*, 1290-1326.

Marsh, H.W., Jayasinghe, U.W. & Bond, N.W. (2008). Improving the peer-review process for grant applications. Reliability, validity, bias, and generalizability. *American Psychologist, 63*(3), 160-168.

Melin, G. & Danell, R. (2006). The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. *Science & Public Policy, 33*(10), 702-712.

Merton, R.K. (1973 [1942]). *The normative structure of science*. In: R.K. Merton (Ed.), *The sociology of science* (pp. 267-278). Chicago, IL: The University of Chicago Press.

Olbrecht, M. & Bornmann, L. (2010). Panel peer review of grant applications: What do we know from research in social psychology on judgment and decision-making in groups? *Research Evaluation, 19*(4), 293-304.

Porter, A.L. & Rossini, F.A. (1985). Peer review of interdisciplinary research proposals. *Science, Technology & Human Values, 10*(3), 33-38.

Sandstrom, U. & Hallsten, M. (2008). Persistent nepotism in peer-review. *Scientometrics, 74*(2), 175-189.

Simonton, D.K. (2006). Scientific status of disciplines, individuals, and ideas: Empirical analyses of the potential impact of theory. *Review of General Psychology, 10*(2), 98-112.

Simonton, D.K. (2008). Scientific talent, training, and performance: Intellect, personality, and genetic endowment. *Review of General Psychology, 12*(1), 28-46.

Smith, S.R. (2001). The social construction of talent: A defence of justice as reciprocity. *Journal of Political Philosophy, 9*(1), 19-37.

Van Arensbergen, P. & Van den Besselaar, P. (2012). The selection of scientific talent in the allocation of research grants. *Higher Education Policy, 25*, 381-405.

Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (forthcoming). The selection of talent as a group process. A literature review on the social dynamics of decision-making in grant panels. *Research Evaluation*.

Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (forthcoming). Different views on scholarly talent: What are the talents we are looking for in science? *Research Evaluation*.

Van Balen, B. (2010). *Op het juiste moment op de juiste plaats. Waarom wetenschappelijk talent een wetenschappelijke carrière volgt*. The Hague: Rathenau Institute.

Van den Besselaar, P. More competition, better science - on the predictive validity of grant selection. In: S. Hinze, et al., (Eds.), *Translational twists and turns: Science as a socio-economic endeavor*; Proceedings 18th STI, Berlin, 2013, 385-392.

Van den Besselaar, P. & Leydesdorff, L. (2007). *Past performance as predictor of successful grant applications*. Den Haag: Rathenau Instituut.

Van den Besselaar, P. & Leydesdorff, L. (2009). Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation, 1*8(4), 273-288.

Van den Besselaar, P. & Van Arensbergen, P. (2013). Talent selection and funding of research. *Higher Education Policy, 26*, 421-427.

Van den Brink, M. (2009). *Behind the scenes of science: Gender in the recruitment and selection of professors in the Netherlands*. Nijmegen: Radboud Universiteit.

Viner, N., P. Powell & Green, R. (2004). Institutionalized biases in the award of research grants: A preliminary analysis revisiting the principle of accumulative advantage. *Research Policy, 33*, 443-454.

Wennerås, C. & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature, 387*, 341-343.

Zinovyeva, N. & Bagues, M. (2011). *Does gender matter for academic promotion? Evidence from a randomized natural experiment?* IZA Discussion Paper No. 5537.

# 5 Gender differences in scientific productivity: a persisting phenomenon?[1]


## Abstract

There is substantial literature on research performance differences between male and female researchers, and its explanation. Using publication records of 843 social scientists, we show that performance differences indeed exist. However, our case study suggests that in the younger generation of researchers these have disappeared. Moreover the top is no longer strongly dominated by men, as the share of top performing women has increased considerably. If this indicates a new trend, a cultural change in developed societies, where women increasingly outperform men in all levels of education, is also becoming effective in the science system.

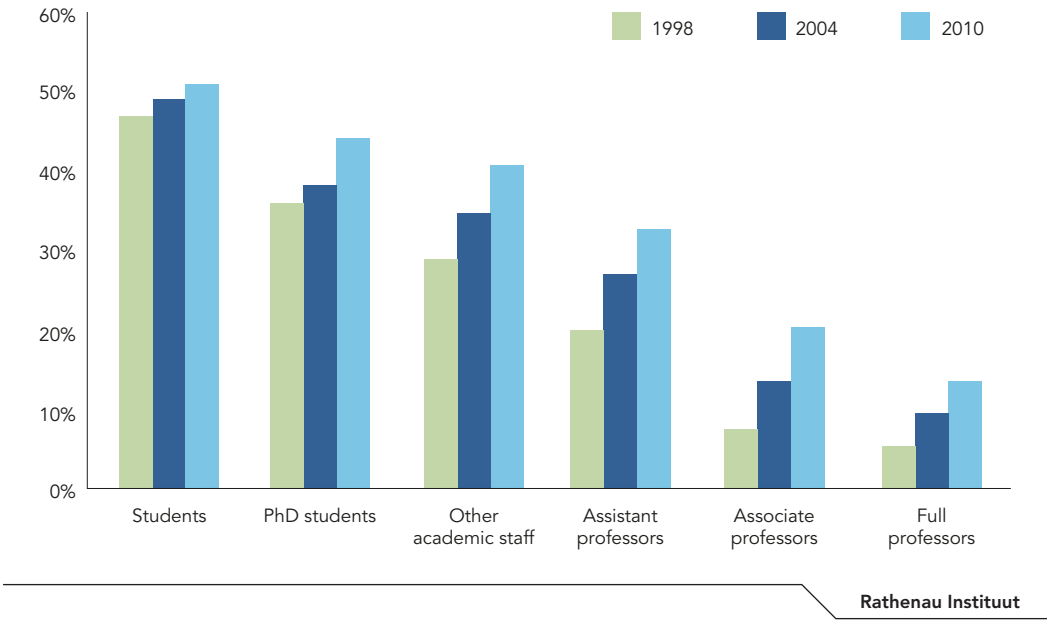

## 5.1 Introduction

The academic world has been dominated by men for a long time. However, the share of women in academia is gradually increasing. Worldwide female students nowadays even outnumber male students, with 55% in the UK and USA and with 59% in the Scandinavian countries (OECD, 2010). And of the new entrance in European higher education about 55% is female.[2] Figure 1 shows the percentage of women in different academic positions in the Netherlands. There, the position of women in higher academic positions is even lower than elsewhere. The growing share of women is characteristic for all positions, however the general rule still is 'the higher the rank in academia, the lower the number of women' (Brouns, 2000; De Weert, 2001; Timmers, Willemsen & Tijdens, 2010). Although female researchers are improving their position, the process is rather slow. Is the weak position due to women having in average fewer ambitions in pursuing an academic career? Are career decisions characterized by gendered social closure, structurally disadvantaging women? Or are women weakly represented in high ranks because their male colleagues outperform them? In this paper we will address the last question by focusing on differences in research performance between male and female researchers.

1 This chapter is a slightly modified version of the paper published earlier in Scientometrics as Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar P. (2012). Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics, 93*(3), 857-868.). For details about the changes see footnote 4 and 12.

2 Of course this differs between the various fields of study. In most science, technology and engineering fields, the share of women is low.

**Figure 1** Share of women in academic positions in the Netherlands 1998-2010 (source: VSNU)




Ample evidence has been provided for a productivity difference between men and women over time, with men producing more research output than women (Abramo, D'Angelo, Caprasecca, 2009; Cole & Zuckerman, 1984; Ledin, Bornmann, Gannon et al., 2007; Long, 1992; Nakhaie, 2002; Penas & Willett, 2006; Prpic, 2002; Symonds, Gemmell, Braisher et al., 2006; Taylor, Fender & Burke 2006; Xie & Shauman, 1998). However, with regard to citations per publication no gender differences were found (Penas & Willett, 2006; Ledin et al., 2007; Tower, Plummer & Ridgewell, 2007), or even a difference in the opposite direction; women having a higher citation score than men (Long, 1992; Powell, Hassan, Dainty et al., 2009). The lower research productivity of women implies that female researchers receive in average a lower total number of citations than men do.

Zuckerman (2001) suggest four different types of explanations of the *productivity puzzle* (Cole & Zuckerman, 1984): scientific ability, self-selection, social selection, and accumulated disadvantage. According to the scientific ability explanation, male and female academics have different biological and psychological characteristics that directly affect the research output. However no direct gender effect has been found in earlier research (e.g. Xie & Shauman, 1998).

The self-selection explanation argues that scientific productivity is influenced by the individual choices of the academics themselves. Several studies confirm the influence of individual choices. For example, women more often interrupt their career to have children and start a family (Prozesky, 2008). Having children causes a decline in research productivity growth, more for women than for men (Fuchs, Von Stebut & Allmendinger, 2001; Hunter & Leahey, 2010). Women were also found to initiate their careers at a later age than men (Karamessini, 2004; Prozesky, 2008). This also holds for their publication career: women produce fewer publications than men

during the first decade of their career, but later in their career they more or less catch up with male researchers (Long, 1992; Symonds et al., 2006). Other factors which are found to affect research productivity and can be considered as self-selection are marital status[3], career ambitions, amount of research time, extent of specialization, discipline, reputation of the university and department, international network (collaboration and co-authoring), and academic rank (Allison & Long, 1990; Bland, Center, Finstad et al., 2006; Carayol & Matt, 2006; Dundar & Lewis, 1998; Leahey, 2006; Lee & Bozeman, 2005; McNamee, Willis & Rotchford, 1990; Prpic, 2002; Puuska, 2010; Taylor et al., 2006). Many of these factors have a gender dimension, as women in average work at lower ranks, in less prestigious institutions, have in average less experience and a weaker (inter)national network. They also specialize less (Leahey, 2006) and more often concentrate on teaching and service, and therefore spend less time on research (Snell, Sorensen, Rodriguez et al., 2009; Taylor et al., 2006). However one should recognize that these factors cannot always be fully ascribed to self-selection. For example, decisions related to collaboration and academic rank are partly in the hands of other people and the organization of the university.

Zuckerman's third type of explanation, social selection, outlines how research productivity of women is affected by gender-based decisions made by others (Zuckerman, 2001). Just as in society in general, there may exist mechanisms of discrimination in the social organization of science (Prpic, 2002). Men outnumber women in positions of formal power, authority and high income (Timmers et al., 2010; Xie & Shauman, 1998). Research on professorial appointments shows there are gender differences in the selection and recruitment procedures. A clear disparity was found in the success rates of male and female applicants to the disadvantage of females (Van den Brink, Brouns & Waslander, 2006). This implies that career decisions are characterized by gendered social closure (Van den Brink, 2009).

A similar situation has been observed in the procedures of grant allocation. Quite some research has focused on gendered aspects of peer review, especially since Wenneras and Wold (1997) published their study on nepotism and sexism in science. They showed that women needed a higher performance to be as successful as male researchers. And, researchers without committee members in their network needed much higher performance than those with an adequate network. A similar study on grant applications in the Netherlands confirmed that gender matters (Brouns, 2000). However it showed that the way it matters varies for different disciplines. Whereas in some disciplines in case of equal average publication scores more men than women were evaluated as excellent, less productive women also obtained grants in other disciplines. Replicating the study of Wenneras and Wold ten years later, Sandstrom and Hallsten (2008) found no sexism anymore; female researchers even had a slightly better chance than males. Clearly, the council studied in both papers changed its policy in the meantime. However, nepotism was as strong as before. If that is the case, this may still influence female researchers, as male researchers generally have better networks than female researchers (Fuchs et al., 2001; Kyvik & Teigen, 1996) and collaboration influences performance (Lee & Bozeman, 2005). Furthermore, women receive less academic support and mentoring than men (Fuchs et al., 2001; Landino & Owen, 1988). This may disadvantage women too, as academic careers depend on support by academic mentors (Van Balen, 2010).

3 Other evidence suggests that the effect of marital status is less univocal (Fox, 2005).

The factors described above may overlap and constitute the source of other events influencing research productivity. For example status in science can be both the cause and effect of scientific collaboration. The same holds for the relation between scientific status and publication productivity (Fox, 2005). The accumulation of decisions or events over time generally placing women at a disadvantage is called cumulative disadvantage (Zuckerman, 2001). However, if productivity differences relate to individual (often gendered) factors, such as ambition, focus on research and changing gender roles and responsibilities in family life (Prozesky, 2008; Taylor et al., 2006; Xie & Shauman, 1998), one may expect that gradually changing gender roles in the last decades may have resulted into changed behavior.

In a recent review, Ceci & Williams (2011) discuss the evidence about discrimination against women in science, in journal reviewing, grant funding, and in hiring. They suggest that no evidence is available that supports the current discrimination against women in science. As a consequence, the unequal position of women in science is based on quality differences between male and female researchers that may partly be based on free choices, and partly on discriminatory arrangements in society at large - e.g., inequalities related to division of domestic work and child care. If this is correct, a careful analysis of these performance differences between male and female researchers is necessary - especially an analysis of changes in performance differences over time. We would actually expect changes, as women increasingly perform better at all levels in the educational system (Buchmann, DiPrete & McDaniel, 2008; Pekkarinen, 2008).

## 5.2 Research question

In this study, we answer the question of whether the gendered productivity differences are persistent or whether they change over time. As it was suggested that the productivity gap occurs in the early career (Symonds et al., 2006), we especially focus on the gendered performance differences among the youngest generation. Research performance in this paper is defined in terms of productivity (number of publications), and in terms of impact (number of citations).

## 5.3 Materials and methods

Comparing male and female researchers requires a good identification of the population. We use data on research grant applications in the Netherlands to analyze productivity differences. The dataset[4] covers about 1100 applications, in a three years period (2003-2005), covering three programs: early (ECG) and advanced career grants (ACG), and an open competition scheme (OC), all within the social sciences.

4 For this chapter, we redid the data collection compared to the original publication (Van Arensbergen et al., 2012). There we used data based on *automatic* coupling of application data with Social Sciences Citation Index data, using family name and first initial. This is of course never perfect, and cannot avoid error based on homonyms and synonyms, applicants using different first initials in applications compared to publications, and titles taken for first names. For this version, we recollected all data *manually* in order to avoid all these problems. We also extended the search, as we now also included non-Dutch addresses and Science Citation Index-expanded publications. In the previous version only publications with a Dutch address and indexed in the Social Sciences Citation Index were included. We now also used a slightly longer citation window (up to 31-12-2006). The new data indeed show many small and several larger differences compared to the original dataset. However, the results of the analyses and the conclusions remain highly similar, suggesting that the error in the data even out. Details about the new data collection will be published elsewhere (Van den Besselaar, forthcoming).

1. The young career grant scheme is meant for researchers who got a PhD within the previous three years. The grant allows them to continue to develop their ideas further.
2. The advanced career scheme is for senior researchers with a long (up to 15 years) post-doctoral experience, and who have shown the ability to successfully develop their own innovative lines of research and to act as coaches for young researchers. The grant allows them to build their own research group.
3. The open competition is for professors and senior researchers. They can apply for a 4-year full-time PhD research project or a 3-year full-time postdoc project.

We had name, field, and institutional affiliation of all applicants. This enabled us to find almost all applicants on the Web, and through that we could identify their résumé and their publications. Those applicants that could not be found were removed from the analysis[5]. This set of applicants can be considered as a good representation of active social science researchers, as active researchers are expected to apply regularly in these programs. As several researchers applied two or more times during the three years, the number of researchers is smaller than the number of applications: 843 researchers, of which 269 (32%) female. The advanced career applicants and the open competition applicants belong to the established generation. The young career grant is clearly for the new generation of scientists. This means we can distinguish two generations of researchers:

1. 355 young researchers, having finished their PhD studies within the last three years;
2. 488 established researchers, generally within the associate or full professor rank.

Full and associate professors are generally older than 40, with an average of 51 years and a standard deviation of 7 years. Those with an ACG grant are on the younger side within this group: they are in average 40 years old with a standard deviation of 4 years.

The ECG grantees represent the young researchers; in our sample, they are between 27 and 41 years, with a few older: researchers who got their PhD at an older age. In average, the young researchers are 33 years old, with a standard deviation of 3 years.

For this paper we define research performance as the number of articles in scholarly (peer reviewed) journals, and as the number of citations received. Research managers and science policy makers increasingly emphasize this type of output and the performance indicators based on it.[6] More specifically, we measured scholarly performance of all researchers, in terms of publications from the *three years before the application* and citations to these publications received until the end of 2006. So we take recent performance and not lifetime performance into account. We used the Web of Science for measuring performance. As we had the résumés and publication lists of almost all applicants, we did not have the problem of synonyms and homonyms (of names). Thus, only the correct publications were included.

5 This also explains some of the differences with the original article.

6 Of course, this does not cover all scientific output, let alone the societal output of researchers (De Jong, Van Arensbergen, Daemen et al., 2011).

The social sciences are heterogeneous, as they consist of psychology, education, pedagogy, anthropology, sociology, communication studies, geography, demography, economics and law. As publication and citation patterns differ between these fields, performance should be standardized in order to use the social and behavioral sciences as one population. However, as the next table shows, three fields dominate the applications: psychology, economics and law. In this paper, we therefore do the analysis first for the (unstandardized) total sample, and then repeat it for the psychology and economics individually.

**Table 1** Applications by field and funding instrument

| | ECG | OC + ACG |
|---|---|---|
| Psychology | 97 | 146 |
| Law | 40 | 110 |
| Economics | 105 | 104 |
| Sociology | 27 | 55 |
| Political science | 13 | 31 |
| Communication | 6 | 15 |
| Geography | 12 | 16 |
| Anthropology | 12 | 9 |
| Education | 50 | 9 |
| Demography | 1 | 5 |
| Grand Total[7] | 355 | 488 |

Rathenau Instituut

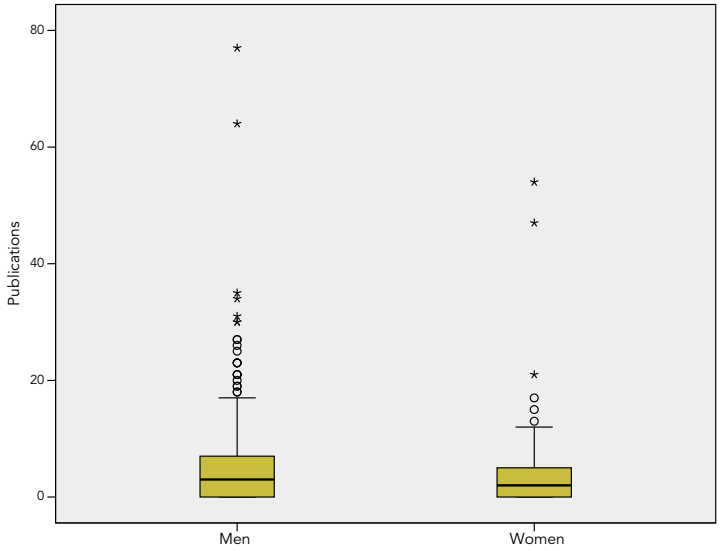
OC = open competition;
ACG = advanced career grant;
ECG = early career grant

## 5.4 Gender differences

First of all, distribution of research performance is heavily skewed. Generally, a small number of the researchers produce the far majority of publications, and a large amount of researchers have a very small output - therefore we use non-parametric statistics. More specifically we use the Monte Carlo method, a powerful simulation technique for obtaining an accurate significance level for highly skewed distributed data, especially regarding small samples.

7 The grand total is smaller than the sum of field counts, as in the overall analysis for applicants who submitted in more than one field only the most recent application was included.

**Figure 2** Productivity by gender, established generation social sciences, 2003-2005

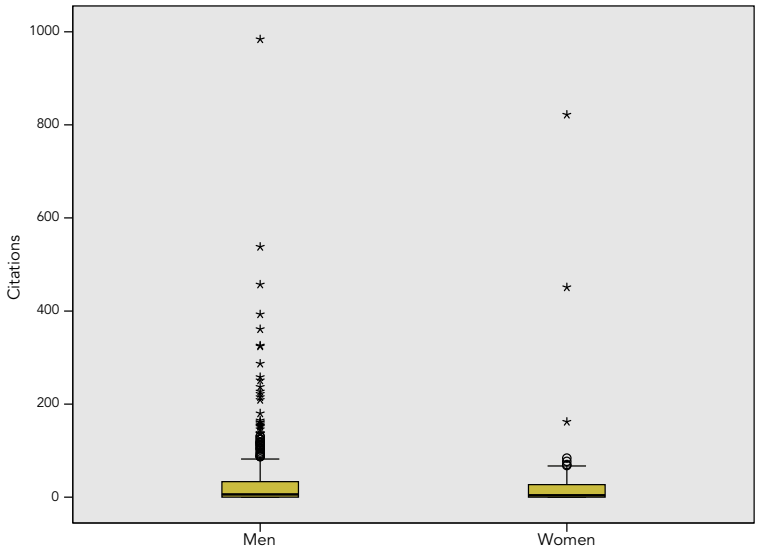



Rathenau Instituut

In the established generation, we have 488 applicants, of which about 22% are female. In the three years period, male researchers did publish in average more than female researchers (mn[8]=5.4 publications versus mn =4.0). The distribution of publications by gender for the established generation (ACG and OC) is shown in figure 2. Clearly, the distributions are very skewed, and we test whether these distributions differ significantly. They do (men mdn[9] = 3 versus women mdn = 2, Mann-Whitney *U*[10] = 18514.5, *p* (Monte Carlo, 1-tailed[11])= 0.048).

8 mn = mean
9 mdn = median
10 From now on will be referred to as *U*
11 From now on will be referred to as *p*

**Figure 3** Impact by gender, established generation social sciences, 2003-2005



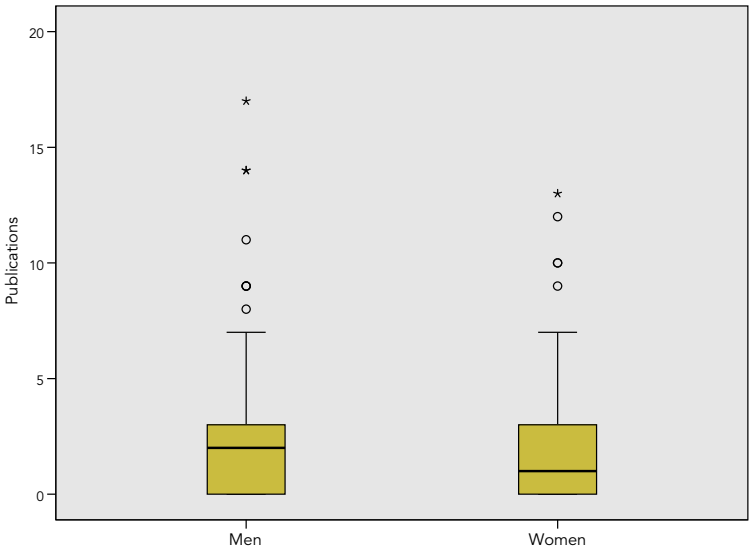

Also in line with earlier findings, in the established generation male researchers receive more citations than female researchers do (mn= 36.6 versus mn= 27.2). The differences are smaller than in the publications. Figure 3 presents the again skewed distributions, which are found to differ (men: mdn= 6 versus women: mdn = 4, $U$=18933.0, $p$=0.092).

## 5.5 Changing gender differences?

We repeated the analysis for the young generation (ECG applicants) with a different result. First of all, of the 355 applicants, about 45% are women. This is a huge increase compared with the established generation (women 22%). In the young generation of scientists, the publication differences have decreased and seem to be disappearing (figure 4). Male researchers still have a slightly higher average number of publications than female researchers (mn=2.3 versus mn=1.9) but the gender difference has become much smaller. That also holds - to a somewhat lesser degree - for the rank order comparison (mdn=2 versus mdn =1, $U$=14202.5, $p$=0.068).

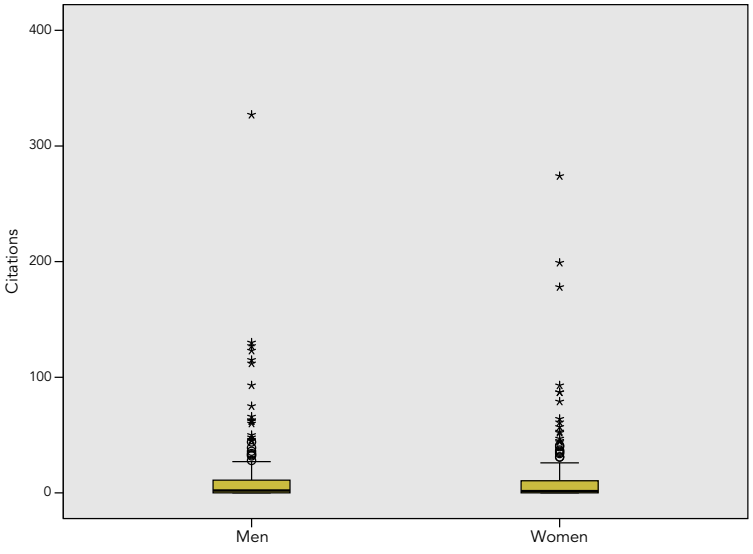
Also the citation patterns have changed, and differences have disappeared more or less (figure 5). Male researchers have a somewhat higher median (mdn=2 versus mdn=1.5) but a lower average (mn=13.4 versus mn=14.0). The Monte Carlo test fails to show a significant difference between the distributions ($U$=15077.0, $p$=0.292).

**Figure 4** Productivity by gender, young generation social sciences, 2003-2005




Rathenau Instituut

**Figure 5** Impact by gender, young generation social sciences, 2003-2005




Rathenau Instituut

When we look specifically at the top of the distribution (the 9-10%), women considerably caught up with men (table 2). While in the established generation 4.6% of the women belonged to the top compared to 10.3% of the men, of the young women 8.8% belongs to the top. In terms of impact the catching up by women is even stronger. In the top 10% impact ranks, women are overrepresented in the young generation. This result differs from what we found for the established generation and is generally found in the literature: an overrepresentation of female researchers in the lower part of the distribution, and an overrepresentation of male researchers in the higher part of the distribution.

**Table 2** Performance by gender

| | Established generation | | Young generation | |
|---|---|---|---|---|
| | % men in top | % women in top | % men in top | % women in top |
| Top ± 10% nr. publications * | 10.3 | 4.6 | 10.8 | 8.8 |
| | | | | |
| Top 10% nr. citations** | 12.1 | 2.8 | 9.2 | 10.6 |

Rathenau Instituut

*: For older generation: >13 publications, for younger generation: >5
**: For older generation: >92 citations, for younger generation: >39

## 5.6 A more detailed view on specific disciplines: psychology and economics

The previous analysis was done at the level of the social sciences as a whole. What if we focus on specific disciplines? We took two social science disciplines with the highest number of applications and in which English language journal articles are the main form of scholarly output.

### 5.6.1 Psychology

Also within the group of the established psychology researchers, men (N=104) in average outperform women (N=42) in publications (mn=12.1 vs mn=7.7; mdn=10 vs mdn=4, $U$=1276.5, $p$=0.000) and in citations (mn= 107.3 vs mn=64.3; mdn=60.5 vs mdn=29, $U$=1352.0, $p$=0.000).

The younger generation (N=97) consists of more women than men, about 55%. Here, the picture is different - in line with the findings for the social sciences as a whole. Output differences between male and female researchers have strongly declined (Mn=4.1 vs Mn=3.1; Mdn=3 vs Mdn=3, $U$=1028.5 $p$=0.158) in the younger generation, as have citation differences (Mn=35.4 vs Mn=24.7; Mdn=16 vs Mdn=9, $U$=1064.0, $p$=0.234).

As shown in table 3, the female researchers are underrepresented in the higher part of the ranking of the established generation. However, in the top of the younger generation ranking they are still underrepresented, although their representation has slightly increased.[12]

12 In the previous version of the paper (Van Arensbergen et al 2012), we found that women outperformed men in the top of the distribution. This has changed with the new data. The potential reason is that we now also include publications in the Science Citation Index, reflecting the more (life) science part of psychology - characterized by higher numbers of publications and citations. Men are overrepresented there. Field normalization might change this again.

**Table 3** Performance by gender - Psychology

| | Established generation | | Young generation | |
|---|---|---|---|---|
| | % men in top | % women in top | % men in top | % women in top |
| Top ±10% nr. publications* | 8.7 | 4.8 | 13.6 | 5.7 |
| Top ±10% nr. citations** | 11.5 | 4.8 | 13.6 | 5.7 |

Rathenau Instituut

*: For older generation: >23 publications, for younger generation: >7
**: For older generation: >224 citations, for younger generation: >66

### 5.6.2 Economics

In line with the general findings, in economics established male researchers have more publications (mn=5.0 vs mn=3.0; mdn=3 vs mdn=3, *U*=376.0, *p*=0.281), and receive more citations (mn=18.2 vs mn=7.7; mdn=9 vs mdn=5, *U*=362.0, *p*=0.221) than established female researchers do. The differences are considerable, however not statistically significant. This could be ascribed to the small number of women (N=9) among the established generation of economists.

In the younger generation, performance differences have changed in favor of women. Average publication and citation numbers are found to be almost equal. More specifically with regard to publications, men do slightly better (men: mn=1.6 vs women: mn=1.2; men: mdn=1 vs women: mdn=1, *U*=904.5, *p*=0.100). In terms of citations, women do slightly better (men: mn=4.6 vs women: mn=5.2; men: mdn=2 vs women: mdn=0, *U*=902.5, *p*=0.091). However, according to the Monte Carlo tests these differences are not significant.

Table 4 shows that in the established generation women are not present in the in the top 10% of the population. Yet, they are to a considerable extent entering the higher performance groups. In the younger generation about 7% of the women economists belongs to the top 10% in terms of publications compared to 13% of the men. When we look at the top performing researchers with regard to number of citations, women are even stronger represented than men. This suggests a similar generational trend as observed within social sciences as a whole. Also in line with the general trend is the increase in the share of women. In the established generation (N=104), women are some 9% and in the younger generation (N=105) this has increased to 27%. However, compared to the share of women among the young researchers within psychology (55%) and the social sciences as a whole (45%) this can still considered to be low.

**Table 4** Performance by gender - economics

| | Established generation | | Young generation | |
|---|---|---|---|---|
| | % men in top | % women in top | % men in top | % women in top |
| Top ±10% nr. publications* | 10.5 | 0 | 13.0 | 7.1 |
| Top ±10% nr. citations** | 10.5 | 0 | 9.1 | 10.7 |

Rathenau Instituut

*: For older generation: >10 publications, for younger generation: >3
**: For older generation: >48 citations, for younger generation: >14

## 5.7 Conclusions and discussion

Our analysis suggests that the gendered performance differences are disappearing. In the older generation, men outperformed women in terms of publications and citations, but this is not any more the case in the younger generation. In other words, the traditional performance differences seem to disappear over time. This is in line with experiences in other parts of the education system, where female pupils and students are increasingly even doing better than male.

This finding is significant as earlier studies found that the performance gap between male and female researchers emerged in the early career phase (Symonds et al., 2006), and exactly in this phase the differences seem to be disappearing. This also suggests that the gendered division of domestic labor, and gender differences in motivation and career planning, may be weakening. As publication and citation scores are increasingly influencing academic careers, the disappearing performance differences may be a stimulus for changing gender relations within science. Of course, the question has to be answered as whether performance differences now emerge in later phases of the research career, a question that requires additional - preferably longitudinal - research.

The current analysis is restricted to the social sciences, and it would be useful to extend the analysis to other fields, such as science, technology, engineering and medicine. Possible performance differences in these fields may be partly due to the low number of female researchers in many of these fields. However, it is also often argued that men have better math and science capacities than women, which would lead to performance differences. This question has been studied intensively, and research suggests these differences - as far as they exist - are decreasing over time (EACEA, 2009; Hyde, Fennema & Lamon, 1990).

Moreover, this study is on a west European case. As the position of women (and consequently of female researchers) differs between countries, the introduction of a cross-cultural perspective would be another useful extension.

Our study indicates that the gender distribution in the group of active social science researchers has changed considerably. In the older generation only about 22% of the applicants are female, in the younger generation this has increased to 45%. Within economics the share of women can still be considered to be low, although it tripled up to 27%. Within psychology, female researchers even have become the majority in the younger generation. If 'mass' explains performance, the remaining performance differences (in fields were the share of women is still relatively low) may disappear when women enter those research fields at a larger share. In those fields, efforts to increase the number of female researchers remain important.

## 5.8 References

Abramo, G., D'Angelo C. A. & Caprasecca, A. (2009). Gender differences in research productivity: A bibliometric analysis of the italian academic system. *Scientometrics, 79*(3): 517-539.

Allison, P. D. & Long, J.S. (1990). Departmental effects on scientific productivity. *American Sociological Review, 55*(4): 469-478.

Bland, C. J., Center, B. A., Finstad, D. A., Risbey K. R., & Staples, J. (2006). The impact of appointment type on the productivity and commitment of full-time faculty in research and doctoral institutions. *The Journal of Higher Education, 77*(1): 89-123.

Bornmann, L., Leydesdorff, L. & Van den Besselaar, P. (2010). A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. *Journal of Informetrics, 4*(3): 211-220.

Brouns, M. (2000). The gendered nature of assessment procedures in scientific research funding: The dutch case. *Higher Education in Europe, 25*: 193-201.

Buchmann, C., DiPrete, T. A. & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*: 319-337.

Carayol, N. & Matt, M. (2006). Individual and collective determinants of academic scientists on productivity. *Information Economics and Policy, 18*: 55-72.

Ceci, S. J. & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. Proceedings of the National Academy of Sciences of the United States of America, *108*(8): 3157-3162.

Cole, J. R. & Zuckerman, H. (1984). The productivity puzzle: Persistence and change in patterns of publication of men and women scientists. Advances in motivation and achievement. P. Maehr and M. W. Steinkmap. *Greenwich, JAI Press. 2*: 217-258.

De Jong, S., Van Arensbergen, P., Daemen, F., Van der Meulen, B., & Van den Besselaar, P. (2011). Evaluation of research in context: An approach and two cases. *Research Evaluation, 20*(1): 61-72.

De Weert, E. (2001). Pressures and prospects facing the academic profession in the Netherlands. *Higher Education, 41*(1-2): 77-101.

Dundar, H. & Lewis, D. R. (1998). Determinants of research productivity in higher education. *Research in Higher Education, 39*(6): 607-630.

EACEA (2009). Gender differences in educational outcomes: Study on the measures taken and the current situation in europe. Brussels, European Commission / Education, Audiovisual and Culture Executive Agency.

Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science, 35*(1): 131-150.

Fuchs, S., Von Stebut, J. & Allmendinger, J. (2001). Gender, science, and scientific organizations in Germany. *Minerva, 39*(2): 175-201.

Hunter, L. A. & Leahey, E. (2010). Parenting and research productivity: New evidence and methods. *Social Studies of Science, 40*(3): 433-451.

Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance - a metaanalysis. *Psychological Bulletin, 107*(2): 139-155.

Karamessini, M. (2004). Women's representation and progression in science careers in greece. Athens, KETHI Research Centre for Gender Equality: 1-26.

Kyvik, S. & Teigen, M. (1996). Child care, research collaboration and gender differences in scientific productivity. *Science, Technology & Human Values, 21*(1): 54-71.

Landino, R. A. & Owen, S. V. (1988). Self-efficacy in university faculty. *Journal of Vocational Behavior, 33*: 1-14.

Leahey, E. (2006). Gender differences in productivity - research specialization as a missing link. *Gender & Society, 20*(6): 754-780.

Ledin, A., Bornmann, L., Gannon, F., & Wallon, G. (2007). A persistent problem - traditional gender roles hold back female scientists. *Embo Reports, 8*(11): 982-987.

Lee, S. & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5): 673-702.

Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces, 71*(1): 159-178.

McNamee, S. J., Willis, C.L. & Rotchford, A.M. (1990). Gender differences in patterns of publication in leading sociology journals, 1960-1985. *The American Sociologist, 21*(2): 99-115.

Nakhaie, M. R. (2002). Gender differences in publication among university professors in canada. *Canadian Review of Sociology and Anthropology-Revue Canadienne De Sociologie Et D Anthropologie, 39*(2): 151-179.

OECD (2010). from http://stats.oecd.org/Index.aspx?DataSetCode=RNENTAGE.

Pekkarinen, T. (2008). Gender differences in educational attainment: Evidence on the role of tracking from a finnish quasi-experiment. *Scandinavian Journal of Economics, 110*(4): 807-825.

Penas, C. S. & Willett, P. (2006). Gender differences in publication and citation counts in librarian ship and information science research. *Journal of Information Science, 32*(5): 480-485.

Powell, A., Hassan, T.M., Dainty, A.R.J. & Carter, C. (2009). Exploring gender differences in construction research: A european perspective. *Construction Management and Economics, 27*(9): 803 - 807

Prozesky, H. (2008). A career-history analysis of gender differences in publication productivity among south african academics. *Science Studies, 21*(2): 47-67.

Prpic, K. (2002). Gender and productivity differentials in science. *Scientometrics, 55*(1): 27-58.

Puuska, H. M. (2010). Effects of scholar's gender and professional position on publishing productivity in different publication types. Analysis of a finnish university. *Scientometrics, 82*(2): 419-437.

Sandstrom, U. & Hallsten, M. (2008). Persistent nepotism in peer-review. *Scientometrics, 74*(2): 175-189.

Snell, C., Sorensen, J., Rodriguez, J. J. & Kuanliang, A. (2009). Gender differences in research productivity among criminal justice and criminology scholars. *Journal of Criminal Justice, 37*(3): 288-295.

Symonds, M. R. E., Gemmell, N. J., Braisher, T. L., Gorringe, K. L. & Elgar, M. A. (2006). Gender differences in publication output: Towards an unbiased metric of research performance. *Plos One, 1*(1): e127.

Taylor, S. W., Fender, B. F. & Burke, K. G. (2006). Unraveling the academic productivity of economists: The opportunity costs of teaching and service. *Southern Economic Journal, 72*(4): 846-859.

Timmers, T. M., Willemsen, T.M. & Tijdens, K.G. (2010). Gender diversity policies in universities: A multi-perspective framework of policy measures. *Higher Education, 59*(6): 719-735.

Tower, G., Plummer, J. & Ridgewell, B. (2007). A multidisciplinary study of gender-based research productivity in the world's best journals. *Journal of Diversity Management, 2*(4): 23-32.

Van Arensbergen, P., Van der Weijden, I. & Van den Besselaar, P. (2012). Gender differences in scientific productivity, a persisting phenonemon? *Scientometrics, 93*: 857-868.

Van Balen, B. (2010). Op het juiste moment op de juiste plaats. Waarom wetenschappelijk talent een wetenschappelijke carrière volgt. Den Haag, Rathenau Instituut.

Van den Besselaar, P. & Leydesdorff, L. (2009). Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation, 18*(4): 273-288.

Van den Brink, M. (2009). Behind the scenes of science: Gender in the recruitment and selection of professors in the Netherlands. Nijmegen, Radboud Universiteit.

Van den Brink, M., Brouns, M. & Waslander, S. (2006). Does excellence have a gender? A national research study on recruitment and gender. *Employee Relations, 28*(523-539).

VSNU. from http://www.vsnu.nl/Universiteiten/Feiten-Cijfers/Personeel/Vrouwen-in-wetenschappelijke-posities.htm.

Wennerås, C. & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature, 387*: 341-343.

Xie, Y. & Shauman, K.A. (1998). Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review, 63*(6): 847-870.

Zuckerman, H. (2001). The careers of men and women scientists: Gender differences in career attainment. Women, science and technology: A reader in feminist science studies. M. Wyer. New York, Routledge: 69-78.

# 6 Determinants of success in academic careers[1]

The competition for top positions in university rankings has put a stronger emphasis on the quality of university staff. Recruitment of excellent scholars is a core activity for university HRM. In this study, we compare the careers of pairs of similar researchers that were considered as very talented in their early careers. Of every pair, one has a continued academic career, whereas the other does not. We investigate to what extent success in academic career is determined by cultural, social and intellectual capital, and organisational and contextual factors.

## 6.1 Introduction

Higher education (HE) and research are increasingly global, as is indicated by the growing obsession with rankings (Deem et al., 2008; Labi, 2008; Mok & Chan, 2008). To reach the academic top, recruiting and keeping the best staff is crucial, as a critical mass of competent highly skilled people is decisive for excellence (Ivancheva & Gourova, 2011). Van den Brink (2009) describes 'recruiting the best scholars' as the core business of universities. As the scientific labour market is increasingly global (Regets, 2007), competition for excellent academic staff is growing (Levin et al., 2006; Mohrman et al., 2008). The reputation of universities plays an important role in attracting excellent researchers (Van Vught, 2008), as does universities' HRM (Thunnissen et al., 2010), and the prevalent career system (Huisman et al., 2002; Van Balen & Van den Besselaar, 2007). The latter lacks transparency, as Van den Brink (2009) has shown, leading among others to an underrepresentation of women in higher positions. If a transparent and formalised method does not exist, what then determines whether excellent talents are preserved for a successful academic career, and do not 'leave the system'?

Empirical studies about academic careers are hardly available. This paper is an explorative and qualitative study of the factors influencing talents to stay in academia. As universities want to select and preserve the best scholars, we focus on careers of high potentials only. Through semi-open interviews, we explore possible relevant factors such as differences in social background (cultural capital), in networks (social capital), in contextual factors (such as the labour market) and in academic performance (intellectual capital).

## 6.2 Research questions

According to Baruch and Hall (2004) the academic career system has unique features, which have made it different from the conventional hierarchical, bureaucratic model of careers. Earlier attempts have been made to describe careers in academia, such as Frost and Taylor (1996), but this was very personal and introspective, with the authors reflecting over their own careers (and thus was past- rather than future-oriented). Research in career development usually concentrates on socio-cognitive factors (De Pater, 2005; Lent et al., 1994) such as the interaction between

1 This chapter has been published as Van Balen, B., Van Arensbergen, P., Van der Weijden, I., & Van den Besselaar, P. (2012). Determinants of success in academic careers. Higher Education Policy, 25, 313-334.

self-efficacy, expectations and career position. Publications on academic career development are mainly restricted to describing potential obstacles for Ph.D. graduates and postdocs entering an academic career or obtaining tenure (Van Balen & Van den Besselaar, 2007), and are less focused on the development of the entire career (Baruch & Hall, 2004).

In a comparison with the academic labour markets of France, Germany and the USA, Musselin (2010) shows that career dynamics differ between these countries. This divergence is mainly caused by the nation-specific 'university configuration'. The three countries differ in terms of the degree of autonomy of the academic profession, the role and frequency of the hiring process, selection principles and incentive mechanisms. For example, Germany is characterised by a strongly hierarchical model and a strong dependence on the external market: to obtain a higher position, one needs to apply to a university in another 'Bundesland'.[2] In contrast, within the USA and to a lesser extent also in France, academics can have a career within the same university. Another example is the procedure of getting tenure. In the US, this is much more formalised than in the other countries. Our study is about career dynamics in the Netherlands. The Dutch case is interesting, as the Netherlands has one of the better performing HE systems, with high publication and citation scores, and a high position in university rankings (THE, 2012).

In this paper we focus on the whole career. Our research question is: Why do some talented researchers have a continued academic career, whereas others do not? More specifically, we will address the following issues:

- Do the scholars who stay report more cultural capital than the leavers, such as higher educated parents and better performance in (pre-) university education?
- Do the talents who stay report a different private situation, especially with respect to childcare?
- Do the stayers report more social capital, such as a better network, more support (mentoring, networking), and access to job and promotion opportunities?
- Do the stayers report more support from the HRM and career system in their university than the leavers?
- Were the labour market conditions better for the researchers who continued their career in university?
- Do the stayers have a higher performance than the leavers in crucial career phases?

## 6.3 Data and methods

This study is based on semi-structured interviews with 42 researchers. The interviews provided us with various types of information. First, we asked them about the relevance of various factors put forward in literature on careers, and that are mentioned in the research questions formulated above. Much of this literature is about countries other than the Netherlands. This study explores whether similar or different mechanisms work in the Dutch HE career system. Second we wanted to be informed about other relevant factors that the researchers experienced themselves to be important for their careers, leading to a description of their career and major events that affected

2 Germany consists of states, in German Bundesland.

their career. In order to analyse the interview material, we organised the data in a timeline with critical career events, such as obtaining Ph.D., receiving important research grants, becoming tenured, being promoted to professor, leaving the university.

The careers of these scholars may be influenced by many different factors, not controlled for in this study. Therefore we decided to use a (case study) strategy of selecting cases with enough variety but also enough similarity. Cases were therefore selected from a variety of disciplines, universities and regions. Within this approach, we selected pairs of a talent who stayed and a talent who left. The pairing is based on similarity in research field and in generation, and may minimalize the uncontrolled effects. This enables us to compare the group of stayers with the leavers, but also between and within the pairs.

In order to create the pairs, we asked HRM departments of universities for excellent 'glad that we could keep them' talents, without indicating their career phase. From the responses, we composed a group of 21 scholars with a thriving university career. In practice these are researchers that were full professors. The stayers were selected in a way to create a variety of discipline, region and gender. In order to find a comparable talent who left, we asked the interviewed stayers to name someone who started an academic career in the same period as they did and who was considered to be very talented, but at some point moved to a non-academic career. Not all stayers could name a leaver, and we could not trace all the people who were named. So we completed the leavers group by asking professors with long-term experience in the same field to provide us with the names of highly talented leavers. A consequence of this recruiting method is that we could not make 21 perfectly matched pairs. Table 1 gives the details.[3]

Apart from the interviews, we collected labour market and performance data. First, data about the academic labour market were obtained from the Netherlands Association of Universities (VSNU).[4] Labour market fluctuations are defined in terms of changes in the number of academic positions within the universities in the period the talents in the research group made their career steps. Due to data availability, this was done on the fairly aggregate level of the main disciplines. In periods the number of relevant positions (e.g., associate professor in the social sciences) has increased, vacancies have been available. In periods of decrease, this was much less the case, taking into account the relatively low mobility in the academic labour market.

3 This sample is not representative for all research careers, as we focus here on the top talents only.
4 http://www.vsnu.nl/Universiteiten/Feiten-Cijfers/Personeel.htm.

**Table 1** The sample distribution according to gender, discipline and region

| | Talents who stayed | | | | | | | | | Talents who left | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Region | West | | North | | East | | South | | Total | West | | North | | East | | South | | Total |
| | M | F | M | F | M | F | M | F | | M | F | M | F | M | F | M | F | |
| Humanities | 2 | 2 | | | | | | | 4 | 1 | 3 | | | | | 1 | | 5 |
| Natural Sciences | 2 | 1 | 1 | | 1 | | 2 | | 7 | 2 | 2 | | | 2 | | 1 | | 7 |
| Social Sciences | 1 | | 1 | 1 | | | | 1 | 4 | 3 | | | 2 | 1 | | | 1 | 7 |
| Technical Sciences | 1 | 1 | | | 1 | | | | 3 | | | | | | | | | 0 |
| Medical Sciences | 2 | 1 | | | | | | | 3 | 1 | 1 | | | | | | | 2 |
| Total | 8 | 5 | 2 | 1 | 2 | 0 | 2 | 1 | 21 | 7 | 6 | 0 | 2 | 3 | 0 | 2 | 1 | 21 |

Rathenau Instituut

Second, we retrieved all the scholars' publications from the Web of Science (WoS) and from Publish or Perish (PoP), in order to determine their academic performance (publications, citations, H-index) in the various phases of their careers. In this way, academic performance of successful and unsuccessful interviewees can be compared. The WoS and PoP data were cleaned: the authors as well as the publications were (manually) disambiguated.

## 6.4 The case

Academic behaviour and career dynamics differ between countries, and this relates to differences between the systems of HE (Musselin, 2010). We therefore briefly describe how the Netherlands' system works.

### 6.4.1 The Dutch system of HE

In the Netherlands, distinction is made between research universities and universities of applied sciences. Research universities offer degree programmes on three levels: bachelor, master and Ph.D. and have the 'lus promovendi', the mandate to award doctorates. All Dutch research universities want a place at the top of university rankings, as can be illustrated by their mission statements.[5] Universities of applied sciences mainly offer bachelor programmes aimed at professional education. Our study focuses at the 14 research universities of the Netherlands, with 40,000 staff and 200,000 students. Ph.D. students are employed by the university on a temporary (4-year) contract to do research and some teaching.[6] Ph.D. students also have to follow courses for their own training and education. Over the last 20 years the number of Ph.D. students successfully defending their thesis in the Netherlands has doubled - from 1898 to 3,736 per academic year (CBS, 2011).

5 http://cf.bc.uva.nl/download/instellingsplan_2007-2010.pdf; http://www.uu.nl/university/utrecht/nl/profielenmissie/hoofdlijnenstrategie/Pages/default.asp; http://www.tue.nl/universiteit/over-de-universiteit/profiel-en-missie.

6 Next to these Dutch universities have the category 'external doctoral students', Ph.D. students not employed by the University. They generally work in other (public) research organisations, in educational jobs or in companies.

Most Ph.D. students leave the university upon graduation, but many aim at an academic career.[7] Therefore they apply for a post-doctoral position (Sonneveld et al., 2010), which is considered as preparation for their first 'real' academic position: assistant professor. The next step could be an associate professor position, succeeded by the final step of becoming a full professor.

Until recently, the numbers of positions at various levels were fixed, and the higher the level, the fewer the positions available. Promotion was dependent on vacancies - not on individual performance. Over the last decade, the career system has become less rigid, and universities are implementing a variety of career systems that are increasingly allowing for promotion trajectories based on individual performance for example tenure track systems (Thunnissen et al., 2010; Van Balen & Van den Besselaar, 2007), leading to a fairly heterogeneous career system.

### 6.4.2 Academic labour market issues

A decade ago the expectation was that Dutch universities would be facing tremendous shortages of eligible candidates for higher academic positions in the near future (Van Vucht Tijssen, 2000). Similarly, the Council for Science and Technology Policy (AWT, 2005) emphasised that more opportunities were needed for research talents to develop their capacities. Others, however, feared an oversupply, leading to a growing gap between the ambitions of young researchers and their chances for an academic career (Hoffius & Surachno, 2006; Keijzer & Gordijn, 2000). In an earlier study we showed that there was neither an under or an oversupply. The real problem is the hierarchical nature of the academic labour force, where it can take a long time for talented young researchers to reach a position as independent researcher: professor (Van Balen & Van den Besselaar, 2007). In this follow-up project, we therefore study the careers of talented researchers, and will try to identify the decisive factors influencing success in academic careers.

### 6.4.3 Criteria for talent?

Although talent is often defined as a natural ability or capacity[8], in an academic context it generally refers to the academic quality of someone's past achievements (Thunnissen et al., 2010; Van Arensbergen & Van den Besselaar, 2012), as emerged by interviewing leading professors in different fields. In their view talented students and researchers produce a very good master's thesis and an excellent doctoral dissertation, and have high grades. Also excellent teaching skills are sometimes mentioned. 'The "talented" students are eager, focused and deeply interested in the discipline, they have passion and drive'. This suggests that criteria for talent relate to research performance, teaching skills and motivation. However, the professors interviewed remained rather vague about the exact criteria used to decide on talent and excellence. They feel that one does not need criteria, as talents will be noticed anyway. This is in sharp contrast with for example, the situation in the US, where tenure depends on explicitly formulated criteria with respect to quality and quantity of research output.[9]

---

7 Ph.D. students are employed by the university, which in practice creates expectations that an academic career is the normal road.

8 http://oxforddictionaries.com/definition/talent; http://www.vandale.nl/vandale/zoekService.do?selectedDictionary=nn&selectedDictionaryName=Nederlands&searchQuery=talent; http://oxforddictionaries.com/definition/talent, last accessed 14 September 2011.

9 For example, http://www.american.edu/provost/academicaffairs/upload/Sociology-Tenure-and-Promotion-Guidelines-FINAL-2-7-2011.pdf.

## 6.5 Findings

When we look at the duration of the various career steps, there are huge differences within and between the pairs. Sometimes it takes only a few years to take the next career step, sometimes many. Table 2 shows the lengths of the phases of several pairs as an illustration. The career data do not indicate that it is necessary to take short career steps to achieve a successful academic career or the other way around. For example, the stayer in pair 7 had a postdoc trajectory of 12 years, but became a full professor in the end, 19 years after obtaining her Ph.D.

**Table 2** Duration of career steps in years

| Talents | Pair 1 | | Pair 2 | | Pair 3 | | Pair 4 | | Pair 5 | | Pair 6 | | Pair 7 | | Pair 8 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S1 | L1 | S2 | L2 | S3 | L3 | S4 | L4 | S5 | L5 | S6 | L6 | S7 | L7 | S8 | L8 |
| Ph.D. trajectory | 5 | 4 | 4 | 5 | 5 | 5 | 4 | 3 | 2 | 4 | 5 | 4 | 5 | 6 | 7 | 13 |
| Postdoc trajectory | 2 | 8 | 4 | 1 | 3 | 4 | 1 | - | 1 | 1 | 6 | 8 | 12 | 5 | - | - |
| Assistant professor | 7 | 1 | 5 | 7 | 7 | - | 7 | 11 | 11 | - | 4 | - | 3 | | 14 | 18 |
| Associate professor | 2 | - | 2 | - | | | 3 | | 3 | | 5 | | 4 | | 5 | - |

S= Talent who stayed; L= Talent who left.

Rathenau Instituut

### 6.5.1 Individual factors

**Cultural capital: Educational level of parents**

Family background, that is parents having undergone HE, used to be an important factor determining the chances for an academic career (Bourdieu, 1988; Van Heek et al., 1958). Among those who influence students' educational expectations, parents play an early and critical role. Wells et al. (2011) present a literature review indicating that the social origin of the family sets the financial, social and cultural context for education. Parents' educational attainment influences their children's educational expectations. Wells et al. (2011) note that the level of education the parents attained indirectly defines the value of HE for their children, but this effect appears to have decreased recently. This last development seems to be confirmed by our study. The interviews showed that the majority of the interviewees' parents did not have an academic degree. This is valid for both groups, the stayers and the leavers. Eight of the interviewees in both groups reported that one parent had undergone HE and 13 indicated that neither parent had. Within the pairs, almost all possible situations arise, except for one: both talents have an academic family background. These findings suggest that the level of education of the parents is not a main factor, a hypothesis that needs further testing.

**Cultural capital: School performance**

Did the successful researchers have better school performances? Most of the interviewees reported high grades during secondary education, which can be seen for the S group as well as the L group (Table 3).

**Table 3** Comparing school success

| | S high grades, L not | Both high grades | Both no high grades | L high grades, S not |
|---|---|---|---|---|
| Secondary education | 3 | 8 | 2 | 4 |
| Master's degree | 4 | 1 | 10 | 2 |
| Ph.D. | 2 | 0 | 7 | 3 |

S= Talent who stayed; L= Talent who left.

Rathenau Instituut

High grades during secondary education motivated several talents to aim for university study. Others did not perform very well during secondary education, and only became motivated later on. The grades did not differentiate between stayers and leavers. Thirteen interviewees of the first group reported high grades during secondary education against 14 of the second group. High grades[10] during secondary education also do not distinguish between the groups at the pair level. Actually, in slightly more pairs, the leavers had higher grades than the stayers, than the other way around, but the differences are small.

During the master's degree, as well as the Ph.D. study period, high grades were less common. For most of the interviewees other stimuli were more important. They mention, for example, interesting research subject, cooperation with their supervisor, extra tasks they performed and participation on student boards. Comparing the pairs shows that at the master's level, the pairs where the stayers scored better were more abundant than those where the leavers scored better. However, at the Ph.D. level, this pattern reversed again. In concluding, educational performance does not seem to differentiate between the two groups of talents. We should emphasise here that the sample is an 'elite' selection. The group that stayed and the group that left were both nominated as 'talent'. All belong to the group of excellent talent. The findings result in the hypothesis that within this group, cultural capital does not seem to influence success in obtaining higher academic positions.

**Family situation**

Several studies have shown that men and women tend to inhabit different sex-based family situations, which may affect development of their academic careers. These include lower marriage rates of women in academe (Probert, 2005), lower geographic and job mobility linked to marriage (Rosenfeld & Jones, 1987), and more significant childcare responsibilities (Hamovitch & Morgenstern, 1977). Furthermore, women who have reached the top in academia seem to be remarkably often childless. Despite improvements in the academic gender balance in recent decades, women are still more likely than men to occupy temporary and part-time positions on a lower level in the academic hierarchy (Baker, 2008). Combining children and an academic career was not easy, as was already noticed in the Netherlands several years ago (Beekes, 1991; Van Doorne-Huiskes, 1979).

10 High grades are defined as 'cum laude' or a comparable level.

More recently, Wolfinger et al. (2008) explored the effects of gender and family formation on academic employment subsequent to receiving a Ph.D. They showed that having a family and children lowers the chance of obtaining tenure-track positions. Single women without young children fare better than their male counterparts on the market for assistant professorship. However, according to Wolfinger et al., family formation cannot account for women's difficulties at later career stages — namely tenure and promotion to full professor. Not all interviewees provided information about their personal situation. Those who did reported that support from a partner is necessary to develop a successful academic career.

> *The partner of S4 did not aspire to have an academic career, although she obtained her PhD. She looked for a job that enabled her to be at home more often and take care of the children. That way she made his career possible.*

The data in Table 4 indicate that the talents who stayed felt more often supported by their partner, than the talents who left. Partners who choose to put all their effort and time in their own career and corresponding residence and working hours, do not stimulate the academic career of the academic talents. As there is no great difference in the number of stayers and leavers having children, this seems to have less influence on the academic career chances. However, the relatively low number of women compared to men who stayed and had children is in line with the observation of Ann Mason (2008) that 'Babies do Matter in Science'. Our findings suggest that this is mainly the case for women.

**Table 4** Family conditions

| | Talents who stayed | | | Talents who left | | |
|---|---|---|---|---|---|---|
| | **Male** | **Female** | **All** | **Male** | **Female** | **All** |
| Have children | 10 | 3 | 13 | 8 | 8 | 16 |
| Supported by partner | 8 | 5 | 13 | 2 | 1 | 3 |
| | **S yes, L no** | | **Both yes** | **Both no** | | **L yes, S no** |
| Have children | 1 | | 10 | 0 | | 2 |
| Supported by partner | 9 | | 3 | 1 | | 0 |

**Rathenau Instituut**

### 6.5.2 Organisational factors

**Social capital: Mentoring**

Sponsorship and mentorship are a 'nurturing process in which a more skilled or more experienced person, serving as a role model, teaches, sponsors, encourages (y) a less skilled or less experienced person for the purpose of promoting the latter's professional and/or personal development' (Anderson & Shannon, 1988, 40). Ehrich et al. (2004) conducted a meta-review of more than 300 research-based articles on mentoring. Their analysis showed that mentoring offers, despite some shortcomings, many far-reaching benefits for mentees as well as for mentors, mentees experience personal support and opportunities for career development. Furthermore

many academic researchers learn from their mentoring relationships how to collaborate and how to interpret social dynamics of collaboration (Mayer et al., 2008). Mentorship can also be important for (pre-doctoral and early-career) research productivity, self-efficacy, grants, and level of promotion and professional network of young researchers (Cameron & Blackburn, 1981; Gardiner et al., 2007; Janasz & Sullivan, 2004; Paglis et al., 2006). A study on the mentoring of junior female academics shows that academics who received mentoring were more likely to stay at the university.

In line with the latter findings, all our interviewees indicated that the support of a mentor, coach or supervisor is very important for an academic career, for some even crucial. 'You will not survive without the support of a mentor'. This can also be reported in a negative sense: some of the interviewees reported that the absence of a coach or supervisor influenced their departure. All four interviewees who had not had a mentor, or even indicated that they were deprived of a mentor, left the university.

S20 met several people in her career who were at some time very stimulating. During secondary education the teachers in Dutch literature were her role models. Literature offered a perspective on the world she could not find at home. During her study at university she was inspired by a teacher in French film studies, later on her PhD supervisor gave her a lot of confidence. When she was a starting scholar she was supported by two female professors, who stimulated her to apply for a full professorship.

**Table 5** Influence of mentors according to the interviewees

| | Talents who stayed | | Talents who left | |
|---|---|---|---|---|
| Felt stimulated by a mentor or sponsor | 17 | | 15 | |
| Career development advices | 14 | | 5 | |
| Did not have a mentor | 0 | | 4 | |
| | **S did, L didn't** | **Both did** | **Both didn't** | **L did, S didn't** |
| Felt stimulated by a mentor or sponsor | 5 | 8 | 0 | 4 |
| Career development advices | 9 | 3 | 4 | 1 |

**Rathenau Instituut**

However, we have to keep in mind that the answers on this question were retrospective views of the interviewees on their career. This view can be influenced by the tendency to attribute 'the failure' of an interrupted academic career to an external cause (Bem, 1972). When comparing the influence of mentors, differences do occur (Table 5 — lower half). This does not hold for the stimulating role of the mentor, but it does for the career advice role. There, in more than 50% of the pairs the talent who stayed owes a piece of crucial career development advice to a mentor, whereas the talent who left did not. And in only a few pairs it worked in the opposite way. These data are in line with findings by other researchers: Mentoring of young scholars is important; giving the right guidance and motivation at the right moment by teaching talented Ph.D. students and postdocs the who, what and how of academia may help (Baruch & Hall, 2004;

Scaffidi & Berman, 2011). But mentoring is certainly not the only factor that counts; also the institutionalised career system matters, which is discussed below.

**Social capital: Networking**

Networking is in many studies described as equally important for an academic career as mentoring or sponsoring (e.g. Shin & Cummings, 2010; Zuckerman, 1991;). Network building very often starts with the mentor or supervisor. The person indicated as mentor was not always, but very often, the same as the Ph.D. supervisor of the talent. The findings from the interviews that network building is important for a successful academic career are in line with earlier research (Burt, 1997, 1998; De Grande et al., 2010).

According to Burt (1997), social capital or networks are crucial for a career. A mentor or sponsor is particular necessary for academics not having social capital themselves. As we showed above, the majority of the talents in our study had no background in academia when starting their academic career. This explains why many of the interviewees refer to the support of a mentor or sponsor.

**Table 6** Influence of networks

| | Talents who stayed | | Talents who left | |
|---|---|---|---|---|
| Owed a job or crucial contact to mentor/sponsor | 17 | | 9 | |
| | S yes, L no | Both yes | Both no | L yes, S no |
| Owed a job or crucial contact to mentor/sponsor | 9 | 5 | 0 | 3 |

Rathenau Instituut

Burt (1998) also argues that professionals starting out in their career should aim at building their social networks within an organisation. On the basis of his research Burt advises 'freshmen' to start a career by borrowing the networks of a mentor or sponsor, but advises talents who are advancing to build their own network with a central position for themselves. This is in line with our findings. We indeed found that successful talents more frequently report that they acquired important relations or job opportunities through their mentor than the leaving talents did - as Table 6 shows.

**University characteristics: Career development system**

Among the factors spontaneously indicated by the interviewees as crucial for their career are the 'career system' and the 'career policies' of the universities. They hold the opinion that the Dutch career system is inflexible and limiting the possibilities for talents. Tenured positions only become vacant when someone else is leaving and this is (with the exception of retirement) not predictable. Universities are reluctant to give the talents a clear perspective. Several Dutch publications (Broersen, 2003; Hoffius & Surachno, 2006) indicate that talents often have to hop from one temporary job to another, awaiting their chance of tenure. Not everyone is able to afford this, since it is difficult to buy a house and support a family with this insecure financial position.

L2 worked for 4 years on a temporary basis, with in total 23 contracts for one university. After some more temporary contracts at another university and two pregnancies, she came to an agreement with the Department on the path to a tenured position. When the end of the temporary contract came, this agreement turned out to be worthless. The Department had financial problems and could not afford to tenure her.

The system does not have enough flexibility to promote and keep those people who have demonstrated their quality. Consequently professors try to keep their talents with vague promises, which they are not able to uphold. The lack of flexibility did turn out to be unfortunate for some of the talents. They reported, for example, that the department or the university was not able or willing to adjust the career rules to people with caring duties. Others, however, mentioned that there were some individual exceptions possible. The rate of flexibility, as reported by the interviewees, depends partly on the faculty and partly on the persuasiveness and effort of individual professors.

L12 agreed with the dean of his Department on the criteria for promotion to associate professor. He should have a number of international publications within three years. After these three years however the rules were changed, he then could get his promotion when he was accepted as a member for a research school. At the moment he met this requirement a new one was added. That was the moment L12 decided to continue his career elsewhere.

**Table 7** Career development system[a]

| | Talents who stayed | | Talents who left | |
|---|---|---|---|---|
| Problem with career system | 5 | | 11 | |
| Problem with HRM practice in the department/university | 8 | | 12 | |
| No problems concerning career development | 12 | | 4 | |
| | **S did, L didn't** | **Both did** | **Both didn't** | **L did, S didn't** |
| No problems concerning career development | 5 | 3 | 9 | 0 |
| Problem with career system | 2 | 3 | 7 | 5 |
| Problem with HRM practice in the department/university | 2 | 6 | 3 | 6 |

Rathenau Instituut

a The interviewees may have experienced problems with the system and with the HRM practice, so these categories do not exclude each other.

The interview reports often paint a portrait of a supervising professor who started optimistically making plans and coaching a promising researcher. However, when leaving the faculty or his or her managerial position, the plans and promises turned out to be worthless. Several interviewed talents who left, but also a few talents who stayed, felt cheated by their university. The career plans made with a supervisor could not be effectuated and rules and standards were changed during the period the talents were trying to meet them. Interestingly, financial issues were not at stake for the talents. None left university for a better salary. Consistency and perspective, that is

knowing that there would still be a position for them in the university in the near future was far more important (see also Thunnissen et al., 2010). In summarising, the talents experienced two kinds of problems in regard to career development:

- Problems related to the career system as such: no tenured positions available at the crucial moment, postdocs hoping and waiting for vacancies.
- Problems related to the HRM practice in the department: unkept promises, no flexibility, no clear career perspectives, no facilities for talents with caring duties (Table 7).

Obviously, talents who left experienced more problems related to academic career developments. These problems concern equally the career system and the HRM practice in the department. These organisational factors indicate what type of talent (HRM) policies could be implemented by universities. We will return to this in the concluding section.

### 6.5.3 Contextual factors: Labour market fluctuations

Some of the talents in this study were born before 1960, and their career suffered from the strong growth of universities in the Netherlands in the 1970s. In the seventies, universities had appointed many staff members to meet the strong increasing student inflow. In the years that followed, these relatively young staff members remained in their tenured positions and no vacancies were available for new talents. Furthermore universities faced severe financial cut-backs, especially in the 1980s. Consequently, for the generation of talents that started their career in the eighties, an academic career was very improbable.

In order to investigate whether labour market fluctuations do play a role in the careers of the talents under study, we used data on the size of the labour force of the universities. For every discipline we obtained the annual change of assistant professors, associate professors and full professor. In some periods the number of staff in a discipline increases. Labour market conditions are more favourable than in periods the number of positions is stable, or in phases with a decline. We distinguished three labour market situations, a growing, a stable and a shrinking labour market, and we related these to the crucial events[11] in the career of the talents.

Data about the labour market fluctuations in the disciplines were available starting from 1990. For some of the pairs we could therefore not analyse the labour market situation at their crucial career events, and the analysis is therefore limited to eight pairs (Table 8). More talents who left had their crucial career events during a shrinking labour market, but the differences are small.

11 That is first tenured position, promotion and departure.

**Table 8** Labour market fluctuations[a]

| Crucial career phase in a | Talents who stayed | | Talents who left | |
|---|---|---|---|---|
| Growing labour market* | 3 | | 3 | |
| Stable labour market | 4 | | 1 | |
| Shrinking labour market | 1 | | 4 | |
| | S did, L didn't | Both did | Both didn't | L did, S didn't |
| Growing labour market* | 2 | 1 | 3 | 2 |
| Stable labour market | 3 | 1 | 4 | |
| Shrinking labour market | | 1 | 4 | 3 |

Rathenau Instituut

a Discipline specific

So far we described several factors that may account for the different career paths of the academic talents. These are factors based on the interviews. Now we will look at their academic performance: are the scholars who stayed at the university simply better than the ones who left? From the university perspective: are universities succeeding in preserving the best academics?

### 6.5.4 Academic performance

In order to belong to the top of the academic world, universities try to attract and to retain the best scholars. Since the number of available job opportunities and positions decreases the higher they are in ranking, researchers leaving the university is unavoidable. However, for most doctorate holders, this is not a voluntary choice (De Grande et al., 2010; Schwabe, 2011). For universities this is not a problem, as long as the best scholars stay. Academic performance should therefore be decisive for the development of one's academic career. According to Long and McGinnis (1993) historical analyses indicate that quantity of publications is the most important factor predicting rank advancement of academic scientists.

We compared the academic performance at various stages of the careers.[12] This enabled us to compare the publication and citation scores at the moment one left with the scores of his or her staying counterpart at the same moment. This informs us whether academic performance determines careers: do better performing researchers stay, and do less performing researchers leave?

12 For several reasons a straightforward performance match was not possible for all 21 pairs, for example, when the person that left did so in an early career phase when academic performance is still modest. For one third of the pairs, a performance comparison turned out to be possible.

**Table 9** Comparing performance of eight pairs

| | Publications | Citations |
|---|---|---|
| *Performance when obtaining Ph.D.* | | |
| Pairs with S perform higher than L | 4 | 5 |
| Pairs with S perform equal to L | 2 | 2 |
| Pairs with S perform lower than L | 2 | 1 |
| *Performance at start tenure (track)* | | |
| Pairs with S perform higher than L | 4 | 4 |
| Pairs with S perform equal to L | 4 | 0 |
| Pairs with S perform lower than L | 0 | 4 |
| *Performance after next promotion/leave* | | |
| Pairs with S perform higher than L | 2 | 3 |
| Pairs with S perform equal to L | 2 | 0 |
| Pairs with S perform lower than L | 4 | 5 |

Rathenau Instituut

S= Talent who stayed; L= Talent who left

In Table 9 we show the performance indicators for three crucial phases in the careers. We include in this analysis eight pairs, four from the sciences, one from the social sciences and three from the humanities. For the first four, we used the WoS for measuring the number of publications and citations, and for the last four we used PoP.[13] The pairs are compared at three moments in their careers: (a) the moment of obtaining their Ph.D. degree; (b) the moment where the researcher obtained a tenured or tenure track position. Of course, more important career events may have taken place, and in that case, we have selected events that are almost similarly positioned in the life cycle: a similar number of years after the Ph.D.; (c) the moment when one of the two leaves, and this compared with a promotion in the career of the other that takes place at about the same moment in terms of 'career age'. For every of the three career events, we distinguish three groups: a group where the stayer performs better than the leaver (S>L); a group where both perform about equal (S¼L); and a group where the leaver performs better than the stayer (SoL). The table shows the size of each of the groups for both performance indicators: (publications and citations) we classified the pairs in three groups. Table 9 shows the resulting distribution over the six categories.

a. At the time of obtaining their Ph.D., the future stayers (S) who outperform the future leavers (L) are the largest group, followed by the equal pairs and the leavers outperforming the stayers.
b. At the mid-career event, this still holds for publications, but not any more for citations: the leavers are as often better than the stayers as the other way around.

13 In both cases, we had to clean the data in order to have the right persons included. Especially in PoP, we unified the publications that appeared in the list in different versions.

c. In the final stage, the picture has changed more radically, and the pairs where leavers outperform stayers have become the majority for publications as well as citations.

These results suggest that there is no systematic relationship between the career success and the commonly used indicators for scholarly performance. Within the group of talented scholars, academic performance does not seem to determine success in a university career.[14]

### 6.5.5 Combining factors

So far we have analysed the data per factor individually. However, interviewees often mention a combination of factors that have affected their career. Moreover the interviews suggest that success is the effect of a number of cascading factors and accumulating advantages, whereas accumulating disadvantages determine whether a talented researcher leaves the university. To test this, we compared the stayers and leavers not in terms of scores on specific factors, but in terms of the number of positive and negative factors they reported. To illustrate this we counted the factors that, in the view of the interviewees, had implications for their academic career. These factors are: (i) support by partner; (ii) a stimulating mentor; (iii) career development advices; (iv) no problems experienced by the academic career system; (v) positive labour market at the crucial career phases.

**Table 10** Pairs and combined effects

| | |
|---|---|
| *S* more positive factors than *L* | 12 |
| Equal number of positive factors | 4 |
| *L* more positive factors than *S* | 1 |

**Rathenau Instituut**

Table 10 indeed shows that in 12 out of 17 pairs, the talented scholar that continued an academic career accumulated more positive factors than the one that left. Only in one pair is it the opposite. In four pairs, both talents reported an equal number of positive factors. Comparing the two groups (in Table 11) supports this finding.

14 In a study comparing successful grant applicants with good rejected applicants, we also found that past performance did not differ between the two groups (Van den Besselaar & Leydesdorff, 2009).

**Table 11** Combined effects

| Balance | Talents who stayed | Talents who left |
|---|---|---|
| 5 factors positive | 2 | 0 |
| 4 factors positive | 5 | 2 |
| 3 factors positive | 8 | 2 |
| 2 factor positive | 6 | 8 |
| 1 factor positive | 0 | 7 |
| 0 factors positive | 0 | 2 |

Rathenau Instituut

## 6.6 Conclusions and discussion

Human resources are recognised as being the key to the creation, commercialisation and diffusion of innovation (Auriol et al., 2010). In this respect, academic scholars are the top of educational hierarchy and specially trained to conduct research and teaching, representing a particularly specialised group in the 'human capital stock' of a society (Schwabe, 2011). The competition for top positions in university rankings has put stronger emphasis on the recruitment of excellent scholars. An important question in HE policy is whether excellent talents are preserved for an academic career, and do not 'leave the system'. In this study we explored which factors influence a successful academic career in the Dutch situation.

The literature analysis and the interviews were performed to see to what extent the factors generally said to influence career development were of importance for the talented scholars and to identify other factors that were important from the viewpoint of the talents themselves. A comparison of researchers who continued and discontinued an academic career seems to confirm the importance of some of the factors (e.g. social capital), whereas others did not (cultural and intellectual capital) differentiate between staying and leaving.

Cultural capital was not found to influence the career paths of stayers and leavers, as no real differences were revealed between stayers and leavers in their educational background and of their parents. Furthermore, support of a partner is necessary in the development of an academic career, while having children was not found to be influential. However, only 50% of the female stayers have children compared to 100% of the males. This suggests that having children does matter for successful female scholars.

The interviews also indicate the importance of social capital. The support of a mentor is important for a successful academic career. In more than half of the pairs the talented stayers owe a piece of crucial career development advice to a mentor whereas the leaver did not. Mentors are also important with regard to access into the networks that they provide, as we found that in more than half of the pairs the talented stayers owe a job or crucial contact to a mentor whereas the leavers did not. Talented scholars reported problems with career policy [organisational factor] as developed on the macro level — within the Netherlands — and with the career policy and opportunities within their own department. In one third of the pairs, the leavers indicate prob-

lems with the career system whereas the stayers did not. These results suggest that the career system does influence career success of talented scholars.

By linking the crucial events in the career of the talented scholars to the labour market conditions [contextual factor] in that period of time, we have shown that talented researchers left academia more often during a shrinking labour market. However, it should be noted that differences are small and because of limited data concerning labour market fluctuations only a proportion of the pairs could be compared. We found no systematic relationship between the career success and the academic performance of highly talented scholars, measured as the number of publications and citations. In the final career phase the leavers even seemed to outperform the stayers, showing that high productive researchers are not always preserved by the university system.

### 6.6.1 Accumulation of (dis)advantages

In conclusion, our exploration does not reveal one deciding factor that determines which talents are preserved for the university. We actually found a wide variety of combinations. Our results suggest that academic careers of talented researchers are stimulated (for those that stayed) or inhibited (for those that left) by an accumulation of advantages or disadvantages, including, according to several of the interviewees, coincidences. Future research on a larger and representative sample of pairs of talented scholars should further test this hypothesis.

What does our study mean for HE career policy? If accumulation of positive and negative factors more than talent as such is decisive, universities could take a proactive stance towards talent management, and create conditions in which competition based on talent and performance is supported, more than only 'being at the right place at the right moment'. The coincidence factor could be more decisive for the Dutch situation than elsewhere because of the absence of a transparent career system and the lack of criteria for career advancement. Unlike for example, the procedures in major US research universities, early career talents often lack information about the number of articles or books they need to have written in order to get tenure. Above that, appointments are often not based on advice of external scholars assessing the work of the candidates (Van den Brink, 2009). Our findings suggest that university policy should aim at clarifying the criteria for career advancement, and at introducing individual performance-based promotion mechanisms. This development has to some extent started with new initiatives, such as mentorship programmes and tenure track systems (Van Balen & Van den Besselaar, 2007). Both could contribute to a consistent, transparent and better 'talent management' approach.

## 6.7 References

Anderson, E.M. & Shannon, A.L. (1988). Toward a conceptualization of mentoring. *Journal of Teaching Education, 39*(1), 38-42.

Auriol, L., Felix, B. & Schaaper, M. (2010). Mapping careers and mobility of doctorate holders. *STI Working paper 2010/1*. Paris: OECD.

AWT (2005). Briefadvies inzake onderzoeksloopbanen. Den Haag: Adviesraad Wetenschaps-en Technologiebeleid, http://www.awt.nl/uploads/files/Briefadviezen/ briefonderzoeksloop-banen.pdf.

Baker, M. (2008). Ambition, confidence and entrepreneurial skills: Gendered patterns in academia. http://www.tasa.org.au/.../Baker,%20Maureen,%20Session%2046%20PDF.pdf.

Baruch, Y. & Hall, D.T. (2004). The academic career: A model for future careers in other sectors? *Journal of Vocational Behavior, 64*(2), 241-262.

Beekes, A. (1991). De hordenloop. Ontwikkelingen in de achterstand van vrouwelijke op mannelijke academici aan Nederlands universiteiten in de periode 1960-1985. Utrecht: ISOR.

Bem, D. (1972). Self-perception Theory, in L. Berkowtiz (ed.) Advances in Experimental Social Psychology. Vol. 6, New York: Academic Press, pp. 1-62.

Bourdieu, P. (1988). Homo Academicus, Cambridge: Polity Press.

Broersen, S. (2003). Werken in de wetenschap: de loopbaanpositie van postdocs, Leiden: Research voor Beleid.

Burt, R.S. (1997). The contingent value of social capital. *Administrative Science Quarterly, 42*(2),339-365.

Burt, R.S. (1998). The gender of social capital. *Rationality and Society, 10*(1), 5-46.

Cameron, S.W. & Blackburn, R.T. (1981). Sponsorship and academic career success. *The Journal of Higher Education, 52*(4). 369-377.

CBS (2011). Wetenschappelijk onderwijs: gepromoveerden aan universiteiten. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=71247ned&D1=0&D2=a&D3=a&D4=a&VW=T, accessed 20 June 2011.

De Grande, H., De Boyser, K. & Van Rossem, R. (2010). Carrièrepaden van doctoraathouders in België: loopbaanpatronen naar wetenschapsgebied. Gent University.

De Pater, I. (2005). Doing things right or doing the right thing. A new perspective on the gender gap in career success. Ph.D. thesis, University of Amsterdam.

Deem, R., Mok, K.H. & Lucas, L. (2008). Transforming higher education in whose image? Exploring the concept of the "world-class" university in Europe and Asia. *Higher Education Policy, 21*(1), 83-97.

Ehrich, L.C., Hansford, B. & Tennent, L. (2004). Formal mentoring programs in education and other professions: A review of the literature. *Educational Administration Quarterly, 40*(4), 518-540.

Frost, P.J.& Taylor, M.S. (1996). Rhythms of Academic Life: Personal Accounts of Careers in Academia, Thousand Oaks: Sage Publications.

Gardiner, M., Tiggemann, M., Kearns, H. & Marshall, K. (2007). Show me the money! An empirical analysis of mentoring outcomes for women in academia. *Higher Education Research and Development, 26*(4), 452-442.

Hamovitch, W. & Morgenstern, R.D. (1977). Children and the productivity of academic women. *Journal of Higher Education, 48*(6), 633-645.

Hoffius, R. & Surachno, S. (2006). Tussen wens en werkelijkheid: carrièreperspectieven van jonge onderzoekers, Leiden: Research voor Beleid.

Huisman, J., de Weert, E. & Bartelse, J. (2002). Academic careers from a European perspective: The declining desirability of the faculty position. *Journal of Higher Education, 73*(1), 141-160.

Ivancheva, L. & Gourova, E. (2011). Challenges for career and mobility of researchers in Europe. *Science and Public Policy, 38*(3), 185-198.

Janasz, S.C. & Sullivan, S.E. (2004). Multiple mentoring in academe: Developing the professorial network. *Journal of Vocational Behavior, 64*(2), 263-283.

Keijzer, B.S.C. & Gordijn, E.H. (2000). Resultaten arbeidsmarktenquête jonge wetenschappers. Den Haag/Utrecht: NWO-ORP/LAIOO.

Labi, A. (2008). Obsession with rankings goes global. *Chronicle of Higher Education*, 17 October.

Lent, S.W., Brown, S.D. & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1),79-122.

Levin, H.M., Jeong, D.W. & Ou, D. (2006). What is a world class university? http:// www.tc.columbia.edu/centers/coce/pdf_files/c12.pdf.

Long, J.S. & McGinnis, R. (1993) Rank advancement in academic careers: Sex differences and the effect of productivity. *American Sociological Review, 58*(5), 703-722.

Mason, A. (2008). 'Babies do matter in science'. Chronicle of Higher Education, 17 October, http://chronicle.com/jobs/news/2008/10/2008101701c.htm.

Mayer, A.P., Files, J.A., Ko, M. & Blair, J.E. (2008). Academic advancement of women in medicine: Do socialized gender differences have a role in mentoring? *Mayo Clinic Proceedings, 83*(2), 204-207.

Mohrman, K., Ma, W. & Baker, D. (2008). The research university in transition: The emerging global model. *Higher Education Policy, 21*(1), 5-27.

Mok, K.H. & Chan, Y. (2008). International benchmarking with the best universities: Policy and practice in mainland China and Taiwan. *Higher Education Policy, 21*(4), 469-486.

Musselin, C. (2010). The Market for Academics. New York: Routledge.

Paglis, L.L., Green, S.G. & Bauer, T.N. (2006). Does adviser mentoring add value? A longitudinal study of mentoring and doctoral student outcomes. *Research in Higher Education, 47*(4), 451-476.

Probert, B. (2005). I just couldn't fit it in: Gender and unequal outcomes in academic careers. Gender, *Work and Organization, 12*(1), 50-72.

Regets, M. (2007). Brain circulation: The complex national effects of high-skilled migration. http://www.oecd.org/dataoecd/59/57/38387715.pdf, last accessed 20 June 2011.

Rosenfeld, R.A. & Jones, J.O. (1987). Patterns and effects of geographic mobility for academic men and women. *Journal of Higher Education, 58*(5), 493-515.

Scaffidi, A.K. & Berman, J.E. (2011). A positive postdoctoral experience is related to quality supervision and career mentoring, collaborations, networking and a nurturing research environment. *Higher Education, 62*(6), 685-698.

Schwabe, M. (2011). The career paths of doctorate graduates in Austria. *European Journal of Education, 46*(1), 153-168.

Shin, J.C. & Cummings, W.K. (2010). Multilevel analysis of academic publishing across disciplines: Research preference, collaboration, and time on research. *Scientometrics Online First*, 6 May, doi 10.1007/s11192-010-0236-2.

Sonneveld, H., Yerkens, M. & Van de Schoot, R. (2010). PhD Trajectories and Labour Market Mobility. Survey of Recent Doctoral Recipients at Four Universities in the Netherlands. Utrecht: IVLOS.

Times Higher Education (2012). World reputation rankings 2012.

Thunnissen, M., Fruytier, B. & Van den Brink, M. (2010). Beleid en beleving. Onderzoek naar jongetalentenbeleid op Nederlandse universiteiten, Nijmegen: Radboud Universiteit & Hogeschool Utrecht.

Van Arensbergen, P. & Van den Besselaar, P. (2012). The selection of scientific talent in the allocation of research grants. *Higher Education Policy, 25*(3), 405-406.

Van Balen, B. & Van den Besselaar, P. (2007). Universitaire onderzoeksloopbanen. Den Haag: Rathenau Instituut.

Van den Besselaar, P. & Leydesdorff, L. (2009). Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation, 18*(4), 273-288.

Van den Brink, M. (2009). Behind the scenes of science. Gender practices in the recruitment and selection of professors in the Netherlands. Ph.D. thesis, Radboud University Nijmegen.

Van Doorne-Huiskes, J. (1979). Vrouwen en beroepsparticipatie. Een onderzoek naar gehuwde vrouwelijke academici. Ph.D. thesis, Utrecht University.

Van Heek, F., Vercruysse, E.W., Veld-Langeveld, H.M., Kuiper, G., Van Braam, A. & Korstanje, B. (1958). Sociale stijging en daling in Nederland, deel I, Leiden: Stenfert Kroese.

Van Vught, F. (2008). Mission diversity and reputation in higher education. *Higher Education Policy, 21*(2), 151-174.

Van Vucht Tijssen, B.E. (2000). Talent voor de toekomst, Toekomst voor Talent. Den Haag: Ministerie van OCW.

Wells, R.S., Seifert, T.A., Padgett, R.D., Park, S. & Umbach, P.D. (2011). Why do more women want to earn a four-year degree? Exploring the effects of gender, social origin, and social capital on educational expectations. *Journal of Higher Education, 82*(1), 1-32.

Wolfinger, N.H., Mason, M.A. & Gouldon, M. (2008). Problems in the pipeline: Gender, marriage, and fertility in the ivory tower. *Journal of Higher Education, 79*(4), 388-405.

Zuckerman, H. (1991). The Careers of Men and Women Scientists, in H. Zuckerman, J.R. Cole & J.T. Bruer (eds.) The Outer Circle. Women in the Scientific Community, New York: W.W.Norton, pp. 27-56.

# 7 Conclusions and general discussion

The past decades can be characterized by a growing awareness of the importance of human capital. Across all industries, the quest for talent has generally invaded society. Academia is no exception to this. Universities have openly made it one of their main goals to attract scientific talent. Human resource management is more and more turning into talent management, with several talent programs being implemented to attract, develop and retain excellent academics. Is talent within academia so scarce that we need this focus and these programs? Although there are differences between disciplines in supply and demand, in general it seems there is no real shortage of talent. The term 'War for Talent' (Michaels, Handfield-Jones and Axelrod, 2001) was introduced in times of increasing demand and decreasing supply, referring to talent as being a highly sought-after commodity. However, currently with regard to science, there rather seems to be a 'war between talents', with an oversupply of talents and an undersupply of academic career opportunities.

In this study we have aimed to unravel the frequently used term 'talent'. What exactly is academic talent? Here a distinction needs to be made between definitions and practices of recognition. Asking members of the academic community to describe the characteristics of talent is, in a way, a bottom-up approach to formulate a definition. We have also used a top-down approach, starting from the outcomes of talent selection processes to determine what is considered to be talent. Because of the growing importance of personal grants in the funding of academics, we chose the Innovational Research Incentives Scheme (*Vernieuwingsimpuls*) of the Dutch Research Council as specific context in considering talent selection. We studied how academic talent is assessed and selected, both by individuals and panels, taking into account the social and competitive nature of these processes. Furthermore, we looked at the role of gender in talent selection and studied performance differences between male and female academics. Finally, we analyzed the careers of academics who once were labeled as talents to identify factors supporting or impeding academic careers. This chapter will summarize the main results, instigating reflection on the current mechanisms of talent selection and its potential consequences for individual academics and academia in broader sense. While doing that, we will refer back to the validity of some of the claims and critique described in the introductory chapter. It concludes with several implications for practice and further research.

## 7.1 Summary of most important results

*Chapter 2* analyzed the notion of talent based on 29 interviews with established researchers involved in the allocation of personal career grants. We compared the general notion of talent that academics value in their daily work to the concrete notion of talent within grant allocation processes. Everyone seems to have ideas about what talent is, but making these ideas explicit is quite a challenge. Generally it is found easier to point out talents amongst people we know directly or indirectly. This relates to the tacit dimension of talent, reflected in popular statements like 'you know it when you see it' and 'everyone will recognize real talent'. When predefining talents as those who stand out and can be considered of great value for a research group, university or science in general, experienced scholars largely agree upon which characteristics to ascribe to talent. First of all, they do not excel on one single dimension, but they possess both

social, professional and predominantly individual capital: most importantly, they are social people, have acquired prestigious grants, are able to work very hard, are ambitious and have published extensively. Seemingly, having outstanding research skills and an extensive publication record do not suffice to be considered an academic talent. Rather these academic skills and expertise are a basic requirement, and it is personal and social characteristics which distinguish real talents from their peers. As academic work is increasingly becoming team work, social skills like communication skills, ability to motivate others and to fit in a group, gain more importance. In the context of grant allocation, the definition of talent can be narrowed down to mainly professional capital: talents have published a lot, including as single author and preferably also without their current promoter and on new topics, they have written a well reasoned proposal in a generally comprehensible way, and they have international working experience.

Moreover, the obtainment of grants (especially personal grants) is considered a particularly important indicator for talent, also outside the concrete context of grant allocation. Personal research grants like ERC's starting grants and NWO's Veni and Vidi grants, provide grantees not only with financial resources, but also with considerable academic prestige due to their high symbolic value. Established scholars indicated the necessity of grants for further career development and generally perceived them as one of the few options for early career researchers to stay within science. The high symbolic value not only has to do with being able to acquire funding (a skill of increasing importance also for early career researchers), but with previous recognition of a person's quality, seeing that assessment criteria not necessarily include actual obtainment of grants, but also receiving positive reviews of grant proposals. The general claim of 'you know it when you see it' might therefore be rephrased to 'you know it because they saw it'.

Which components of talent are valued the strongest depend on the context in which evaluation takes place. With regard to talent selection within the Dutch funding program, the Innovational Research Incentives Scheme, the goal is set in advance in the call: to select talented, creative researchers who engage in innovative research. The procedure determines the input, the information available to reviewers. A written résumé and research proposal expose different characteristics compared to a face- to-face interview with the applicant. The interview phase was found to be very important (as also shown in *chapter 4*), enabling panellists to assess especially the candidates' individual and social capital and to test the tenability of their previous reviews based on the written applications.

With regard to the process of grant allocation, panellists indicated they faced two main problems: the broadness of the set of applications they needed to review and the small quality differences within the main share of applicants. As with the panel deliberation, no clear differentiation could be made using the objective and most common criteria (e.g. scholarly performance, international experience, innovativeness of proposal); more tacit and subjective criteria (e.g. personality, fit in the group, perseverance) come into play. Talent cannot fully be captured by objective criteria or expressed in scores even to two decimals places. Our study illustrates this by panellists describing how they reach their final decisions. At the end of the panel meeting when all proposals and applicants were elaborately discussed and scored, a ranking is calculated based on these scores. Several panellists described how they checked this ranking for correctness. Not whether the ranking was calculated correctly, but whether candidate A rightfully stood above

candidate B, disregarding the better score of A. Besides these more subjective criteria, selection outcomes are affected by factors inherent to the social nature of the group decision-making process (e.g. panel composition, speaking order and atmosphere within panel). Many panellists indicated to be aware of these various 'uncontrollable' influences. Despite the thorough selection procedure and time investments, they acknowledged their final decisions regarding the large set of applicants of equal average quality to partly be the product of arbitrary and subjective influences.

More insight on the social nature of grant allocation was detailed in *chapter 3*, which covers an extensive literature review on panel review processes. Many of the review and selection processes within science involve panel reviewing, a group activity. As this entails social interaction, group dynamics influence the review process. For this reason we combined studies from the sociology of science and science policy studies with literature from social psychology. Where the first mainly focus on how peer review affects review outcomes, the latter focuses on actual review processes. Research within social psychology predominantly deals with central mechanisms involved in decision-making processes and the context in which these are carried out. We designed a conceptual model of grant allocation processes, including factors related to the applicants, their application and their network, to the social process of reviewing and to the context in which it is performed. Characteristics of the applicant which were found to affect review outcomes are, for example, past performance, status of previous employers, research trail and research field, more specifically cognitive similarity between applicant and reviewer. With regard to panel deliberation, for example, the composition of the panel affects the represented expertise and sharing of information, the social identities of and status differences between panellists, and the communication within the panel. The decision-making process furthermore is open to influences from factors like group norms, motivational differences and strategic behavior. Finally, the available time and budget, procedural guidelines and the condition of accountability were found as influential context factors. The complexity of these types of processes reduce the predictability of its outcomes. Due to the (social) nature of this process, these influences can never be completely excluded (Lamont, 2009). According to Opthof and Wilde (2009), politics and personal bias affect review outcomes due to the difficulty of their task: while they succeed in filtering out work that should not be granted, it is generally impossible for reviewers to distinguish between good or excellent applications. This chapter explained how grant allocation is the product of human interaction, and therefore liable to a certain extent of subjectivity and arbitrariness.

*Chapter 4* described a quantitative study on grant allocation and the evidence of talent. Statistical analyses of the review scores of about 900 grant applications were conducted. These involved personal career grants from the Innovational Research Incentives Scheme for researchers in different career phases. First of all, we found no clear 'boundaries of excellence', as there were very small differences in scores between those who were successful and those who were just unsuccessful. We did not find any evidence for the general claim (as also often heard in the interview study of *chapter 2*) that 'top talent is easily recognized by everyone'. Furthermore, the moderate correlations between the evaluation criteria suggests that talent is multidimensional and people can excel on different dimensions. Consensus on quality was found to be lower for researchers who were further in their career, as also within the Social Sciences and Humanities compared to Technical and Life Sciences.

The considerable changes in the assessment scores through the various phases of the selection process, indicate that talent evaluation is dependent on how it is organized, confirming that talent is context dependent as shown in *chapter 2*. The interview phase was found to be very influential proven by the considerable changes in the ranking of applicants after the interview, indicating the importance of the new information added to the selection process. At the same time we saw that the weight of reports of external referees, peers who are asked to review an application based on their specific expertise, is modest. Panellists, who cannot be considered real peers to all applications they have to review, can overrule these external reviews, e.g. when they disagree or when the reviews lack a clear motivation (see *chapter 2*). Furthermore, the panel has additional autonomy in decision-making. Finally, no evidence was found for a structural gender bias due to the panel composition. However, our findings do show that when women are better represented amongst the applicants and the panellists, biases in either direction seem to diminish.

Despite the rise of female academics, men are still overrepresented at the higher positions within academia. Can this be related to gender differences in academic performance? *Chapter 5* investigated whether the generally claimed performance differences between men and women still persisted in the younger generation of researchers. Based on publication and citation records of about 845 social scientists, recent scholarly performance was compared between established and early career researchers, in all social science fields together, and separately for psychology and economics. In the established generation, men were indeed found to outperform women. However, in the younger generation these differences had disappeared. Furthermore, in the younger generation the share of women in the top increased considerably compared to the established generation. An important factor that may contribute to the changing differences is the considerable increase of the share of women. Compared to the established generation, the share of women in the younger generation has doubled (from 23% to 45%) and in psychology, women are even overrepresented (55%). If these differences do not emerge in later career phases after all, the disappearing performance differences may be a stimulus for changing gender relations in science, as numbers of publications and citations increasingly influence academic careers, especially in the early career phase.

*Chapter 6* explored which factors influence academic talents to stay in academia or to leave. This qualitative study was based on 42 interviews, labour market data and performance data. We compared the careers of pairs of similar researchers (in terms of generation and research field) that were considered as very talented in their early careers. One had a continued academic career, the other left academia. First of all, we found considerable differences between academic career structures, implying it is not necessary to take short career steps to achieve a successful academic career. Next, whereas cultural capital (educational background of parents and school performance) does not seem to influence career paths, social capital does. Mentor support and network building opportunities were found to be especially important. With regard to family situation, support of a partner is necessary, and as amongst the stayers, far fewer women than men had children, our findings suggest that having children does matter for successful female scholars. Furthermore, an important problem related to the career system that was generally mentioned was the lack of flexibility and clear career perspectives. With regard to intellectual capital (scholarly performance), no systematic differences were found between stayers and leavers. This study does not confirm that the university system always preserves the highly

productive researchers, as leavers were even found to outperform the stayers in the final career phase. Finally, we did not find one single factor determining which talents are preserved for the university, but rather an accumulation of (dis)advantages, including coincidences.

Next, we will reflect on potential consequences of the current selection practices for the supply and development of talent. Taking the results of this study as a starting point, we will outline a few potential scenarios that will refine some of the claims described in the introductory chapter. These scenarios are no direct results of this study, but are meant to stimulate further debate on talent selection and academic career development.

## 7.2 Further reflection upon the current selection system and its accompanying criticism

Science is growing more and more competitive, especially with regard to resources and positions. The general principle of competition is that it should improve quality and transparency. At the higher levels of science and universities this might largely be valid, but at the individual level only a small group benefits from it. Those who 'win' receive considerable recognition and a positive boost to their career (Bloch, Graversen and Pedersen, 2014). But many of those who 'lose' (even when they are considered to be 'just as good') leave academia or stay in temporary positions without any clear career prospects. After years of insecurity and job-hopping, a switch to a career outside academia could be complicated by a high degree of specialization and a higher age (Broersen, 2003). But there seems to be a growing awareness that at higher levels, competition does not only yield positive results. Most extreme are people caught out after unethical behavior and fraud, claiming they felt forced towards this because of severe competition. Starting from our findings and taking them a step further for reflection, what potential consequences do the current selection mechanisms have for the supply and development of academic talent?

### 7.2.1 The significance of publication records and competition

The strongest criticism concerning current scientific practices is directed at the overpowering role of publishing in 'surviving' and building a career within science. Quality nowadays seems to a large extent to be defined as productivity. Universities seem to have internalized the performance culture and rhetoric to such an extent that academics even define and regulate themselves in terms of dominant performance indicators like numbers of publications, citations, or the H-index. Academics are evaluated on their contribution to the performance of the organization instead of their contribution to the professional and intellectual community (Morley, 2005). Consequently, publishing scientific articles seems to have become the goal of academic labour. The enormous pressure to publish is criticized to a growing extent, as it is considered to have detrimental effects, e.g. encouragement of fraud and salami science. In fact in my opinion it leads to an inflation of the scientific article: who still has the time and interest to read the multitude of publications? What is the added value of each publication, more often describing smaller parts of research, to knowledge production or to its societal use? Due to the perception of publishing as a goal of academic work instead of a means to distribute knowledge, less attention is paid to what is really done with newly accumulated scientific knowledge.

High numbers of publications and citations indeed seem to be particularly important to aim for when you want to succeed in academia as an early career researcher. Having an extensive

publication record adds to career opportunities within science, although this study showed that it is a condition but not a guarantee for success. Moreover, our results question the overpowering significance assigned to these performance measures in the debate, as they were not found to be entirely decisive. If they would have been, we would have seen clear cut decisions, but we did not. Where previous studies already showed that allocation outcomes are not just based on publication records (see also Bornmann, Wallon and Ledin, 2008; Melin and Danell, 2006; Van den Besselaar and Leydesdorff 2007, 2009; Hornbostel et al., 2009), this study described what other types of information are used in selection processes. Various personal and social characteristics are included, particularly in face-to-face assessment during the interview, which was found to be highly decisive. This can be considered a positive finding, as academics are involved in more activities than only publishing excellent papers, e.g. teaching, providing supervision, managing research groups, collaboration or distributing knowledge to society. Referring to the last activity, attention is being paid to the contribution of academic research to wider society to an increasing degree.

Nevertheless, publication counts are dominant in the rhetoric of current debates on performance measures and talent selection. This rhetoric seems to be very powerful, although its exact power needs further research. It would indeed become dangerous if the various players within academia will start to believe and adjust to it, as we already see happening. Should this rhetoric affect academic career opportunities, it might induce a constriction of academic talent. Henceforth talent might mainly stand for proven academic achievements instead of potential quality and performance, predominantly supporting career opportunities for academics who manage to publish a lot and obtain lots of external funding. This scenario could potentially lead to a more uniform model of 'the talented academic', as the variety in types of talented academics and their skills would be diminished.

A more uniform model of talent may also be a consequence of growing competition. As science is getting more competitive, especially with regard to research funding, talent selection is also strongly characterized by competition where excellence is required. Generally candidates are compared to each other and the number of candidates who can be identified and selected as talent is limited. The challenge of having to compare candidates is that it forces reviewers to make their criteria explicit. On the one hand this contributes to the fairness of the procedure, as it avoids candidates being assessed on different (implicit) criteria. On the other hand, it causes the 'objective' performance criteria to predominate over criteria related to personal or social skills. Numbers of publications, citations and grants, or years of international experience, are easier to compare than, for example, motivation, perseverance or communication skills (see also Musselin, 2002). While academic talent ideally is perceived as someone excelling on a broad range of characteristics, academic, personal and social, talent selected within these competitive programs might to a large extent be narrowed down to more easily and objectively measurable characteristics, like academic performance. However, when 'measuring' academic performance by counting numbers of publications and citations, one should realize that more need not always be better, as high performing scholars may be less influential than low performing scholars (Waltman, Van Eck and Wouters, 2013).

### 7.2.2 The high symbolic value of grants

Besides publications, research grants are strongly related to prestige. Academics who managed to obtain significant grants are seen as very talented, considerably increasing their chances of receiving subsequent resources compared to those who have not obtained grants (yet), in line with the well- known *Matthew effect*. Universities currently tend to strongly focus on supporting and improving the grant success of their academic staff. There have even been opportunities in the recent past to apply for grants which enabled academics for some period to prepare a grant application for the Dutch national or the European research council. Grant acquirement is often explicitly included in universities' selection criteria in recruitment or promotion procedures.[1]

However, this study showed there is quite some uncertainty regarding grant allocation decisions. Although the few at the top and at the bottom are rightfully granted and rejected respectively, within a large pool of applicants quality differences seem minimal. Moreover, due to the social character of the panel review process, decisions are liable to subjectivity and arbitrariness. Then there is the time and information available to reviewers to found their decision with regard to talent selection. There was a general awareness amongst panel members that the grants they distributed were vitally important financial resources for early career researchers in order to develop an academic career. As these grants therefore are considered *career* grants instead of just *research* grants, the assessment of the applicant is an important part of the selection procedure. The information available on which to base the assessment, is composed of written résumés, a half-hour presentation by a selection of the applicants and perhaps some anecdotal information. Compared to the richness of information that can be gathered by people working close to the applicant, for example, this input can be considered fairly restricted, when we take into account that the resumes of early career researchers are generally not clearly distinctive. Despite acknowledgement of the various uncertainties characterizing the allocation process, these grants are still ascribed major symbolic value. Minor differences in quality are enlarged to major differences in recognition and thus in career opportunities.

### 7.2.3 Gendered impediments

With regard to gender, the rhetorical emphasis on publishing and acquiring funding can be considered to be detrimental to female academics. It reinforces the 'masculinity of excellence', as it forces people to work long days, fulltime and overtime, with inflexible work schemes, to always be available and because it attributes more value to research than to teaching achievements. As women generally are entrusted with more caring tasks, therefore being less flexible and as they are more involved in teaching, women are disadvantaged (Bleijenbergh, Benschop and Vennix, 2013; Mason, Wolfinger and Goulden, 2013). This may result in fewer women making a career in science, encouraging them to seek an alternative career path and consequently resulting in a potential loss of talent for academia.

Next, implications for practice will be given to prevent a loss of academic talent and support academic careers of early career researchers.

1 See for example http://www.ugent.be/nl/actueel/vacatures/zap/erc-consolidator-grants

## 7.3 Implications for practice to uphold and stimulate talent

### 7.3.1 Create more clarity on career perspectives

A complicated issue with regard to talent selection is the question of what is the right moment for selection. While research career opportunities are scarce, it may be more efficient to select at an early stage of people's careers. Resources can be invested more effectively in a smaller group of high potentials, enabling them to fully focus on developing themselves as excellent researchers. Those who do not make it through the selection can concentrate on other career paths, saving many of them years of drudging work, submitting numerous grant applications and cherishing hopes without yielding any career certainty. However, the disadvantage of early selection is that talent has not yet had the opportunity to fully develop, leading to more differentiation between academics. Selecting too early in the development process, when diversity is limited, creates the risk of missing out on talent and potential excellence in the long run.

### 7.3.2 Pay attention to a broad scope of skills

Because of the high level of uncertainty and the lack of clear career perspectives particularly for most of the early career researchers, high potentials should be well supported and supervised. Currently the focus in HRM policy seems to be on organizational interests, increasing the distance from the needs of individual employees (Thunnissen and Fruytier, 2014). Instead HRM programs could stimulate academics to broaden their view and to reflect upon their capabilities, not only in terms of performance, but in terms of skills and experience. Currently, the résumés of researchers predominantly convey their publications, grants, international experience, awards and teaching experience. When this is perceived to be most relevant, other experiences and skills (e.g. related to organizational expertise, entrepreneurship activities or software use) are often overlooked and not described on résumés. Consequently, these other features cannot be reviewed and will not play an important role in selection processes. In order to enable reviewers to broaden their review criteria and to facilitate comparison of these broader criteria, academics should develop portfolios in which these are included.[2] This may enhance their self-awareness of their usefulness and employability inside and outside academia.

Moreover, the state of academic talent may benefit from a softening performance culture. Because of the emphasis on publications and grants, other valuable skills are often far less appreciated. Although criticism on the dominance of productivity measures and requirements is augmenting, no real alternatives are currently implemented. For example, to reduce the publication pressure, academics may no longer be assessed on their total number of publications, but on a limited number of their best achievements (not necessarily scientific publications). To an important extent it is up to academics themselves to initiate change, since they are mainly responsible for the assessment, selection and recruitment within science. It is the academic elite sitting in review panels for the evaluation agencies or research councils who set the norms, criteria and priorities according to which academic activities, and thus academic reward and

2 Part of the ACUMEN project (funded by the European Committee in the 7th Framework Programme) is development of portfolio's reflecting researchers' careers and experiences. These portfolio's enable researchers to provide adequate evidence regarding their various activities and development (http://research-acumen.eu/).

careers, are evaluated (Musselin, 2013; Thunnissen and Fruytier, 2014). Reviewers tend to search for candidates who look like themselves, have similar characteristics, work in similar ways and can be imagined as their colleague; also described as a bias towards homothetic recruitments (Musselin, 2002). This is in line with the rationale of composing grant panels in the Dutch *Vernieuwingsimpuls*: they generally consist out of former laureates. Academics previously identified as talent are equipped with the task of identifying new talent. This automatically sets a kind of blueprint for the candidates they are looking for. Due to the growing importance of review practices and its use by university managers as management tool, the academic elite is reinforced (Musselin, 2013) and therefore the main target group when it comes to changing performance culture or selection criteria.

### 7.3.3 Dare to commit

As shown in this study (*chapter 6*), linking early career academics to a mentor is an important instrument for career support. Furthermore, to prevent loss of talent, high potentials should be offered clear career prospects. Nowadays career opportunities are still strongly dependent on (unpredictable) vacancies and obtainment of grants. Universities should create more possibilities for departments to retain early career academics who are considered to be highly talented. Their talent policy does not need to depend to a great extent on decisions taken by grant panels. If they are convinced of someone's talent, they should have the courage of their convictions and organize their resources to commit themselves to this person and to invest in him or her. In order to rightfully appreciate talent, especially with regard to early career researchers, researchers should not only be assessed within competition, but to a larger extent on their own (unique) merits, offering more room to a broad variety of talent dimensions, including the often distinctive tacit dimension.

Another way to retain talent, particularly female talent, is to enhance the flexibility of the HRM system. Inflexibility is an important factor impeding academic careers, as shown in this study. Career development could for example be supported by more flexibility in working schemes, working on and off site, opportunities to work part time and to better balance professional and family life (Dikkers, Van Engen and Vinkenburg, 2010). This could subsequently stimulate gender equality within academia by contributing to a change in the organizational culture at universities, making it less masculine.

### 7.3.4 Aim for differentiation

To conclude, the current focus on research excellence contributes to the image that academia is only looking for top researchers. Obviously, Dutch universities are not exclusively involved in research, as their main task still remains education. There is more to academic work than conducting research, and not all starting academics aspire to become a professor or a top researcher, (which is just as well as there are insufficient posts for these aspirants). Due to the increased size of the higher education system, universities therefore need to aim for differentiation in academic functions and careers. Besides researchers and teachers, universities need people with expertise in management, organization, fundraising and valorization for example. Excellence can also be applied to these other tasks, and enhancement of academic prestige should also be possible through achievements related to these tasks. Universities should provide room for top teachers or top 'valorizers', who are hardly or not involved in research activities. Various types of career paths

should be stimulated and supported, in which the variety of academic tasks are more equally valued, instead of predominantly valuing research activities.

With regard to evaluation and selection of academics, differentiation of assessment procedures could result in a broader variety of talent being recognized in practice. For example, including interviews in evaluation procedures enables reviewers to assess communication and personal skills, which cannot properly be assessed from written applications or résumés. Variation in procedures prevents a bias towards preferences for a restricted set of skills. Finally, attention should also be paid to diversity within evaluation panels in terms of gender, expertise and status, as panel composition affects the evaluation outcomes.

## 7.4 Implications for future research

In the interviews with panellists gender was not put forward as an important or sensitive issue. While talking about the review process and describing talent, gender was hardly ever mentioned, neither by male nor female respondents. Only to the concluding question, if they had any questions or final additions, several of them responded with 'I expected you to ask me about gender too', followed by 'it did not play any role within our panel'. From other studies we know gender biases can be unconsciously influencing people's behavior, for example, as it is often implicitly embedded in organizational cultures. Asking people about their behavior implies asking them to reconstruct their behavior. This is expected to yield different results from actually observing their behavior. Systematic observations of review or recruitment panels would be very valuable in order to identify the more subtle impact of gender on review processes. They will also enable researchers to better understand the implicit mechanisms which enter into panel review processes and to understand how group dynamics affect the panel decisions. We looked at grant allocation from a social psychological perspective and identified various factors and mechanisms inherent to the social nature of panel reviewing that influence these selection processes. Selection outcomes were found not to be fully determined by track records, leaving them partly unpredictable and intangible. We consider our study a valuable contribution to the existing body of knowledge on peer review and grant allocation, but it requires further investigation on how these factors and implicit mechanisms more precisely affect the actual selection outcomes. Observational studies would be an important and essential method to answer these questions.

This study showed how grants are given strong symbolic value and are perceived to be almost a requirement for academics to develop their academic careers. At the same time we pointed out the power of rhetoric and its potential influence on people's behavior and on policy. Further research is needed on the effects of rhetoric and to find out what the real value and impact of these grants are on individuals' careers. To what extent are they requirements or guarantees for successful academic careers? Is it possible to have a successful career without these grants and does being awarded a prestigious grant always mean you will have a successful career? As career opportunities differ per discipline, comparisons between disciplines would be relevant to get more insight into the impact of grants.

Furthermore, gender differences in scientific performance were found to be disappearing. While in the established generation of researchers men were outperforming women, in the younger generation women performed at least equally to men. This study covered the social sciences.

Although fields like psychology, economics and education are covered rather well in publication databases of Web of Science, fields like law and political science are included to a smaller extent, therefore limiting our results. To determine how solid and generally valid our findings are, replication of this study in other disciplines (with high Web of Science coverage) is recommended. Finally, longitudinal research is needed to study if these performance difference have really disappeared in the long run or if they have shifted towards later career phases.

## 7.5 References

Bleijenbergh, I., Benschop, Y. & Vennix, J. (2013). Group Model Building to support gender equality change in: Achterbergh, J., Benschop, Y., Hendriks, P. & Van de Ven, A. (Eds.): *Op zoek naar het andere: een liber amicorum voor Hans Doorewaard.* Boom: Meppel: 80-96.

Bloch, C., Graversen, E. K. & Pedersen, H. S. (2014). Competitive research grants and their impact on career performance. *Minerva, 52*, 77-96.

Brink, M. van den (2009). *Behind the scenes of science. Gender practices in the recruitment and selection of professors in the Netherlands.* PhD thesis, Radboud Universiteit Nijmegen.

Broersen, S. (2003). *Werken in de wetenschap: de loopbaanpositie van postdocs.* Leiden: Research voor Beleid.

Dikkers, J., Van Engen, M. & Vinkenburg, C. (2010). Flexible work: ambitious parents' recipe for career success in the Netherlands. *Career Development International, 15*(6), 562-582.

Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment.* Cambridge: Harvard University Press.

Mason, M.A., Wolfinger, N.H. & Goulden, M. (2013). *Do babies matter? Gender and family in the ivory tower.* Rutgers University Press.

Michaels, E., Handfield-Jones, H. & Axelrod, B. (2001). *The war for talent.* Boston, MA: Harvard Business Review Press.

Morley, L. (2005). *The Micropolitics of Quality.* Critical Quarterly, 47 (1-2), 83-95.

Musselin, C. (2002). *Academic labour markets: how they work.* Münster conference paper.

Musselin, C. (2013). How peer review empowers the academic profession and university managers: Changes in the relationships between the state, universities and the professoriate. *Research Policy, 42*, 1165-1173.

Opthof, T., & Wilde, A. A. M. (2009). The Hirsch-index: A simple, new tool for the assessment of scientific output of individual scientists: The case of Dutch professors in clinical cardiology. *Netherlands Heart Journal, 17*(4), 145-154.

Thunnissen, M. & Fruytier, B. (2014). Het mobiliseren van Human Capital: een overzicht van 25 jaar HRM-beleid op Nederlandse universiteiten. *Tijdschrift voor HRM, 1*, 1-24.

Waltman, L., Van Eck, N.J. & Wouters, P. (2013). Counting publications and citations: Is more always better? *Journal of Informetrics, 7*, 635-641.

# Nederlandse samenvatting

# Talent Proof: Selectieprocessen in onderzoeksfinanciering en -loopbanen

## Focus op talent en excellente prestaties

Binnen de academische wereld is er een sterke focus op talent en excellente prestaties ontstaan, met name in de eerste fases van de academische carrière. Universiteiten richten zich meer dan ooit expliciet op het aantrekken en behouden van talent door het aanbieden van coaching- en stimuleringsprogramma's en door het werven van nieuwe medewerkers volgens het *Tenure Track*-beginsel. Dit is een loopbaantraject voor *high potentials* dat leidt naar hoogleraarschap als men voldoet aan de vooraf opgestelde prestatie eisen. Voor met name startende wetenschappers zijn talentsubsidies zoals de Veni-, Vidi- en Vici-beurzen (onderdeel van NWO's financieringsprogramma de Vernieuwingsimpuls) en de ERC Starting Grant van groot belang om hun academische carrière te ontwikkelen. Deze Nederlandse en Europese persoonsgebonden subsidies bieden wetenschappers gedurende drie tot vijf jaar de mogelijkheid om hun eigen onderzoekslijn verder te ontwikkelen en ondersteunen hen bij het opzetten van hun eigen onderzoeksgroep.

Het groeiende aantal aanvragen en het lage honoreringspercentage (11-17%), wijzen erop dat dit systeem van talentselectie onder druk staat. Daarnaast klinkt er vanuit de academische gemeenschap in toenemende mate commentaar op de grote publicatiedruk, ook wel beschreven als de *publish or perish*-cultuur. Excellente prestaties staan in de wetenschap voornamelijk voor hoge aantallen publicaties, ook wanneer het gaat om onderzoekers in de beginfase van hun carrière. Tussen deze onderzoekers bestaat er een groeiende competitie om financiering en (het verkrijgen en behouden van) posities. Deze toenemende druk en ontwikkelingen vragen erom het huidige systeem van talentselectie onder de loep te nemen.

## Inzicht creëren in proces van talentselectie

Aan de hand van diverse kwalitatieve en kwantitatieve onderzoeksmethoden poogt deze dissertatie meer inzicht te creëren in het proces van talentselectie. Gezien het grote belang van individuele onderzoeksfinanciering, gaat dit onderzoek specifiek in op selectieprocessen bij de verdeling van persoonlijke onderzoekssubsidies.

De hoofdvragen die hierbij centraal staan, zijn: wat is academisch talent en hoe wordt het geselecteerd? Deelvragen met betrekking tot het selectieproces, zijn: welke mogelijke groepsfactoren zijn van invloed op het selectieproces? Welke rol spelen competitie en publicaties voor academische carrières? Welke genderverschillen kunnen er worden waargenomen als het gaat om academische prestaties? En welke factoren belemmeren of bevorderen de ontwikkeling van academische loopbanen?

## Conclusies

**Verschillende opvattingen over academisch talent**
Wat is academisch talent en hoe vindt talentselectie plaats? Om antwoord te krijgen op deze vragen zijn 29 gevestigde wetenschappers geïnterviewd over twee verschillende onderwerpen: talentselectie aan de universiteit in hun eigen vakgroep en talentselectie binnen het financieringsprogramma de Vernieuwingsimpuls waarbij zij lid zijn van de beoordelingscommissie. Om het selectieproces beter te begrijpen, zijn daarnaast ook de cijfermatige beoordelingen van ongeveer 900 Vernieuwingsimpulsaanvragen geanalyseerd.

Wanneer je ervaren wetenschappers vraagt om academisch talent te omschrijven, zeggen ze vrijwel allemaal: 'De echte talenten herken je meteen.' Toch blijkt het expliciet benoemen van de kenmerken die hen tot talent maken, nog een lastige opdracht. De basisvoorwaarden waarover de talenten beschikken, zijn volgens de geïnterviewden in elk geval: excellente onderzoeksvaardigheden en een uitgebreide publicatielijst. Vervolgens zijn het de persoonlijke en sociale kwaliteiten die de echte talenten onderscheiden van hun vakgenoten. Een belangrijke indicator van talent voor de wetenschappers is daarnaast de verwerving van een prestigieuze onderzoekssubsidie, zoals een Veni-, Vidi- of ERC-beurs. Gezien de lage honoreringspercentages moeten de laureaten wel toptalenten zijn, zo luidt de redenering.

Binnen het selectieproces van de Vernieuwingsimpuls wordt een smallere opvatting over talent gehanteerd. Hierbij ligt de nadruk allereerst op academische prestaties. Uit de interviews en de cijfermatige analyse blijkt wel dat het zwaartepunt gedurende de selectieprocedure verschuift. Aan het begin van het selectieproces is een talent een aanvrager 'op papier' die vooral veel gepubliceerd heeft, ook als eerste auteur, in vooraanstaande tijdschriften en het liefst ook onafhankelijk van zijn of haar promotor; hij of zij heeft beurzen verworven en buitenlandse werkervaring opgedaan, en heeft een goed uitgewerkt voorstel geschreven in algemeen begrijpelijke taal. Later in de procedure, wanneer de commissieleden de aanvragers in levenden lijve kunnen beoordelen, onderscheiden de echte talenten zich sterker op basis van hun enthousiasme, overtuigingskracht en doorzettingsvermogen. Dit betekent dat sommige aanvragers die op basis van hun voorstel en curriculum vitae als echt talent werden beschouwd, na de interviewronde terugzakken naar de middenmoot en vice versa.

Kortom, talent wordt contextueel gedefinieerd vanuit een bepaald referentiepunt, dat verschuift gedurende het beoordelingsproces.

**Talentselectie: omgaan met dilemma's**
De gevolgde procedure is van invloed op de uitkomst van het selectieproces. De meeste geïnterviewde commissieleden binnen de Vernieuwingsimpuls ervaren twee problemen die het beoordelingswerk bemoeilijken: de breedte van de set aanvragen en de kleine verschillen in kwaliteit binnen een grote groep aanvragers.

Een commissie moet een groot gebied dekken in termen van expertise om alle aanvragen te kunnen beoordelen. Dit betekent dat commissieleden slechts voor een klein aantal aanvragen als specialist beschouwd kunnen worden en voor de meerderheid als 'generalist'. Ze beslissen over

voorstellen die buiten hun directe expertise vallen. In deze gevallen waarbij ze de voorstellen minder goed op de specifieke inhoud kunnen beoordelen (bijvoorbeeld het innovatieve karakter of de onderzoeksmethoden), baseren ze zich in de eerste fase sterker op meer algemene criteria zoals het curriculum vitae van de aanvrager en de leesbaarheid van het voorstel.

Naast een klein aantal uitschieters aan de boven- en onderkant, bestaat er een grote groep aanvragers waarbij geen duidelijke differentiatie mogelijk is op basis van de algemene criteria, zoals publicaties, beurzen, en internationale ervaring. Toch moet er binnen deze groep een selectie worden gemaakt. Hiervoor worden additionele criteria toegepast die meer impliciet en subjectief zijn, zoals persoonlijkheid, doorzettingsvermogen en ambitie. De druk om tot een beslissing te komen terwijl duidelijke kwaliteitsverschillen ontbreken, maakt het selectieproces vatbaar voor de invloed van oncontroleerbare factoren zoals subjectiviteit, toeval en groepsdynamiek. Commissieleden zijn zich hier terdege van bewust en zetten dan ook vraagtekens bij een gedeelte van de selectie-uitkomsten. Niet omdat de aanvragers die een beurs toegekend krijgen niet van hoge kwaliteit zijn, maar omdat een gedeelte van de afgewezen aanvragers in kwaliteit nauwelijks van de geselecteerde onderzoekers onderscheiden kan worden.

**Talentselectie: groepsdynamische factoren dragen bij aan besluitvorming**
De subjectieve en veelal oncontroleerbare invloeden op het selectieproces, zoals hierboven beschreven, zijn onvermijdelijk. Niet alleen vanwege de geringe kwaliteitsverschillen, maar ook vanwege het beoordelingsproces zelf, dat zich kenmerkt door sociale interactie en groepsbesluitvorming. De literatuurstudie in dit onderzoek heeft geleid tot een conceptueel model dat diverse factoren onderscheidt die van invloed zijn op het selectieproces in de beoordeling van onderzoeksaanvragen.

Naast factoren gerelateerd aan de aanvragers, hun netwerk, hun voorstel en de selectieprocedure, wordt groepsbesluitvorming beïnvloed door factoren inherent aan sociale interactie. Het model maakt hierbij onderscheid tussen factoren die te maken hebben met de samenstelling van het panel, met de motivatie en inspanning van panelleden en met de verzameling en verspreiding van informatie.

Factoren die een rol spelen ten aanzien van de aanvragers en die van invloed zijn op hun beoordelingen, zijn bijvoorbeeld bewezen prestaties, status van voormalige werkgevers en onderzoeksveld. Belangrijk hierbij is de mate van cognitieve gelijkheid tussen de aanvrager en de beoordelaar. Als er sprake is van een disciplinaire match, heeft de beoordelaar vaak meer expertise en affiniteit met betrekking tot het onderzoek van de aanvrager, wat kan leiden tot een positievere beoordeling. Daarentegen kan een ‘match’ er ook aan bijdragen dat de beoordelaar juist kritischer naar het voorstel kijkt.

Daarnaast beïnvloeden procedurele factoren zoals tijd- en aanvraagdruk en richtlijnen de uitvoering van het beoordelingsproces. In situaties van hoge tijdsdruk besteden groepsleden bijvoorbeeld meer aandacht aan gedeelde dan aan unieke informatie en is er minder ruimte voor alternatieve beslissingen. Eerder onderzoek wijst bovendien uit dat de mate waarin beoordelaars hun keuzes moeten verantwoorden, effect heeft op het groepsproces, met name op de motivatie van de beoordelaars en op het delen van informatie.

Een belangrijke factor in groepsbesluitvorming is de samenstelling van de commissie, aangezien hierdoor wordt bepaald welke expertise beschikbaar is binnen de groep. Welke informatie het zwaarste weegt en de wijze waarop commissieleden dit delen en waarderen, wordt onder meer bepaald door (waargenomen) statusverschillen en onderlinge machtsverdelingen. Ook groepsnormen, motivatieverschillen en de sociale cohesie tussen groepsleden zijn van invloed op het besluitvormingsproces. De exacte wijze waarop veel van deze sociaal-psychologische factoren de uitkomsten van deze groepsprocessen beïnvloeden (bijvoorbeeld de uiteindelijke honoreringen), verdient nader empirisch onderzoek.

**Afnemende genderverschillen in wetenschappelijke productiviteit**
Ondanks de wereldwijde toename van het aantal vrouwen op alle niveaus in de wetenschap, bestaat er in de hogere wetenschappelijke posities nog altijd een sterke ondervertegenwoordiging van vrouwen. Vooral Nederland scoort laag met een gemiddeld percentage van nog geen 15 procent vrouwelijke hoogleraren. De ondervertegenwoordiging van vrouwen is onderwerp van veel onderzoek en heeft geleid tot verschillende verklaringen: mannen en vrouwen zouden verschillen in hun ambities; mannen worden impliciet bevooroordeeld in selectieprocessen; mannen presteren beter dan vrouwen. Aansluitend bij de algemene opvatting dat academische prestaties de loopbaankansen van met name jonge onderzoekers vergroten, neemt deze dissertatie prestatieverschillen als uitgangspunt en stelt de vraag of deze prestatieverschillen nog steeds gelden voor de nieuwe generatie onderzoekers.

Om tot een antwoord te komen op deze vraag zijn de aantallen publicaties en citaties van ongeveer 850 sociaal wetenschappers aan het begin van hun carrière voor drie groepen verzameld en geanalyseerd: voor de sociale wetenschappen als geheel, en specifiek voor psychologie en voor economie. Daarbij is onderscheid gemaakt tussen de gevestigde generatie (Open Competitie- en Vici-aanvragers) en de jonge generatie (Veni-aanvragers).

De vergelijking maakt duidelijk dat de mannen in de gevestigde generatie inderdaad beter presteren dan vrouwen, in de zin dat ze meer publiceren en vaker geciteerd worden. Deze traditionele prestatieverschillen lijken echter in de meer recente periode te verdwijnen. In de jongere generatie worden namelijk geen of aanzienlijk kleinere prestatieverschillen aangetroffen tussen mannen en vrouwen. Ook het aandeel vrouwen in de top van meest publicerende en geciteerde onderzoekers, is in de jonge generatie aanzienlijk toegenomen ten opzichte van de gevestigde generatie.

De resultaten van deze studie zijn met name interessant omdat eerdere studies aantoonden dat de prestatiekloof tussen mannen en vrouwen in de vroege carrièrefase ontstond. Dit is precies de fase waarin de prestatieverschillen nu verdwenen lijken te zijn. Aangezien academische prestaties in termen van publicaties en citaties een belangrijke rol spelen in de loopbanen van - met name startende - onderzoekers, zou het verdwijnen van prestatieverschillen wel eens aanleiding kunnen zijn voor verandering van genderrelaties binnen de wetenschap.

**Duidelijke factoren voor talentbehoud in de wetenschap ontbreken**
Talentselectie draagt er in belangrijke mate aan bij dat wetenschappers hun wetenschappelijke carrière verder kunnen ontwikkelen. Toch is daarmee niet gezegd dat de eens geselecteerde

talenten binnen de wetenschap blijven functioneren. Om meer inzicht te krijgen in factoren die ertoe leiden dat talenten in de wetenschap blijven of hieruit vertrekken, zijn 42 personen geïnterviewd die aan het begin van hun wetenschappelijke carrière als talent zijn aangemerkt. Deze 42 onderzoekers zijn voor de analyse in paren van twee verdeeld - gelijke onderzoekers, in termen van generatie en onderzoeksveld - waarbij een van hen nog binnen de wetenschap werkt en de ander een baan buiten de wetenschap heeft gevonden. Daarnaast zijn ook hun curriculum vitae, hun prestatiegegevens en arbeidsmarktgegevens geanalyseerd.

Snelle carrièrestappen blijken niet noodzakelijk om een succesvolle loopbaan binnen de wetenschap te hebben. 'Sociaal kapitaal' is daarentegen wel van groot belang: ondersteuning van een mentor, mogelijkheden om te netwerken en steun van de partner. Het hebben van kinderen lijkt geen belemmerende factor te zijn, al is het aantal mannelijke blijvers met kinderen aanzienlijk hoger dan het aantal vrouwelijke blijvers.

Belangrijke belemmeringen voor talenten om door te gaan in de wetenschap, zijn de beperkte academische loopbaanmogelijkheden en de inflexibiliteit van het academische carrièresysteem. Vertrekkers wijzen veelal op het ontbreken van heldere carrièreperspectieven, niet nagekomen loopbaanafspraken - waarbij individuele hoogleraren een belangrijke rol spelen - en het niet willen of kunnen aanpassen door de universiteit van prestatie-afspraken, bijvoorbeeld in het geval van mensen met zorgtaken.

Een vergelijking van de academische prestaties binnen de paren laat niet zien dat het altijd de best presterende onderzoeker is die binnen de wetenschap blijft. In termen van aantallen publicaties en citaties is er geen relatie tussen academische prestatie en loopbaansucces. Er zijn zelfs een aantal paren waarbij de vertrekker later in de carrière beter presteert dan de blijver.

Samenvattend blijkt er niet één doorslaggevende factor te zijn die bepaalt of talenten binnen de wetenschap blijven of een baan daarbuiten vinden. Het gaat doorgaans om een opeenstapeling van voor- of nadelen die een academische loopbaan respectievelijk bevorderen of belemmeren. Toeval is hierbij een niet te onderschatten factor. Een volgende carrièrestap was volgens meerdere wetenschappers het resultaat van 'op het juiste moment op de juiste plaats zijn'.

**Vragen bij talent en talentselectie**

In de huidige selectiesystemen binnen de wetenschap wordt veel waarde toegekend aan het verwerven van onderzoekssubsidies. Niet alleen vanwege hun financiële waarde, maar vooral vanwege hun symbolische waarde: het is een belangrijke indicator van talent. Gezien de sterke competitie tussen onderzoekers, moeten de laureaten wel toptalenten zijn. Op basis van de onderzoeksresultaten in deze dissertatie, kunnen echter vraagtekens worden geplaatst bij de sterke symbolische waarde die de academische gemeenschap toekent aan dit type persoonlijke onderzoeksfinanciering.

In de eerste plaats kent het selectieproces meerdere onzekerheden. Door de kleine kwaliteitsverschillen tussen de aanvragers, worden de uitkomsten beïnvloed door factoren die losstaan van de kwaliteit van de aanvragers. De samenstelling van de commissie en de organisatie van de procedure spelen een belangrijke rol. Daarnaast zijn er factoren inherent aan de sociale aard van

het beoordelingsproces, die bijdragen aan een belangrijke mate van onvoorspelbaarheid en toevalligheid in de beslissingen van de commissie. Kleine verschillen in kwaliteit worden door de honoreringsbeslissingen omgezet in grote verschillen in erkenning en daarmee in academische loopbaanmogelijkheden. Dit betekent niet per definitie dat het selectieproces niet goed wordt uitgevoerd. Deze invloeden zijn nu eenmaal niet uit te sluiten in processen waarbij het gaat om mensenwerk. Waar wel iets aan gedaan kan worden, is de wijze waarop met de selectie-uitkomsten wordt omgegaan en de betekenis die hieraan wordt toegekend. Gezien de onzekerheden in het selectieproces zou een bescheidener symbolische waarde van deze beurzen voor verdere loopbaanmogelijkheden gepast zijn.

Wanneer het gaat over excellente wetenschappers en academisch talent, staan naast persoonlijke onderzoeksbeurzen voornamelijk publicaties centraal. In het huidige debat binnen de wetenschap is veel kritiek op de hoge publicatiedruk en de eenzijdige focus op aantallen artikelen als criteria voor kwaliteit. Deze dissertatie laat zien dat publicaties inderdaad van groot belang zijn voor academische loopbanen. Zonder een mooie publicatielijst kun je het binnen de wetenschap wel vergeten. Ze dienen echter als ondergrens waaraan wetenschappers minimaal moeten voldoen, waarna verdere selectie plaatsvindt op basis van andere criteria, zoals ambitie, doorzettingsvermogen en communicatievaardigheden. De huidige focus op publicaties is echter eenzijdig. Publicaties zijn geen goed middel om te differentiëren tussen talenten. Die eenzijdige focus leidt tot een onderwaardering van andere belangrijke academische activiteiten en vaardigheden, zoals onderwijs, begeleiding, leiderschap, management, ondernemerschap en valorisatie. Om academici te stimuleren zich verder te ontwikkelen op deze andere waardevolle gebieden, is het van belang hier ook expliciet waardering aan toe te kennen binnen evaluaties en beoordelingen.

Ten slotte staat deze eenzijdige beoordeling waarbij de nadruk ligt op publicaties en onderzoeksbeurzen, momenteel onder druk. Niet alleen door het huidige debat dat wetenschappers onderling voeren, maar ook door bijvoorbeeld de recente invoering van maatschappelijke relevantie als expliciet criterium in het nieuwe Standaard Evaluatie Protocol (2015-2021). Al het onderzoek dat wordt uitgevoerd aan de Nederlandse universiteiten en binnen de NWO- en KNAW-instituten, wordt volgens dit protocol geëvalueerd. Deze ontwikkeling vereist een verandering in het denken en doen van wetenschappers en bestuurders. Toekomstig onderzoek moet uitwijzen wat de gevolgen van deze ontwikkeling zijn voor de (h)erkenning van talent en de ontwikkeling van academische loopbanen. Worden andere talenten geselecteerd en leidt het tot meer differentiatie van academische functies en loopbanen? En wat zijn de gevolgen voor loopbaanmogelijkheden en ontwikkelingen met betrekking tot gender en diversiteit?

# Dankwoord

Al was het schrijven van dit boekje een behoorlijk solitaire bezigheid, ik had dit nooit kunnen doen zonder de hulp van velen. De afgelopen jaren heb ik het geluk gehad veel inspirerende, boeiende, kritische, hulpvaardige en mooie mensen te ontmoeten, die me op de een of andere manier geholpen hebben mijn onderzoek te (blijven) doen, mijn ideeën vorm te geven en mijn resultaten te delen.

Als eerste wil ik graag Peter en Inge bedanken. Peter, je was een zeer enthousiaste, inspirerende en creatieve promotor. Gesprekken met jou brachten altijd veel nieuwe ideeën en motivatie. Jij hebt me als voormalig afdelingshoofd van SciSA de ruimte geboden mijn onderzoeksinteresse uit te werken en me het vertrouwen gegeven te gaan waar ik voor sta. Van jou heb ik ook een goed inkijkje gekregen in het spel dat wetenschap heet en heb ik geleerd me hierin staande te houden. Inge, je was een geweldige co-promotor voor wie ik veel bewondering heb. Al ben ik erop vooruitgegaan, jouw plan- en organisatietalent zal ik nooit evenaren. Dank voor de fijne samenwerking en voor de steun op de momenten dat ik het nodig had.

Dank gaat ook uit naar mijn commissie, die tijd heeft vrijgemaakt om mijn manuscript kritisch te lezen. Ik bewonder jullie expertise en werk, en voel me vereerd dat jullie mijn werk wilden beoordelen en mij hierover bevragen.

Iedereen die bereid was mee te werken aan een interview, dank voor jullie openheid en interessante verhalen. Zonder jullie was het niet mogelijk geweest mijn onderzoek uit te voeren. Dit geldt ook voor hen die me de mogelijkheid boden in hun commissie mee te kijken hoe discussies over onderzoeksvoorstellen nu in de praktijk verlopen. Dank voor jullie medewerking en vertrouwen.

Het Rathenau Instituut wil ik graag bedanken voor de mogelijkheden die het me geboden heeft om mezelf te ontwikkelen – op meerdere gebieden tegelijk – en om een deel van mijn werk uiteindelijk te publiceren als proefschrift. Het was een zeer leerzame werkomgeving met veel geweldige collega's en inmiddels ook nieuwe vrienden. Dank aan alle collega's voor het kleur geven aan het instituut en de fijne werksfeer. Barend, we waren het niet altijd met elkaar eens, maar dat hield me scherp en vergrootte mijn motivatie om mijn proefschrift af te ronden. Ik heb veel geleerd van onze discussies en ik ben blij dat ik ook kon bijdragen aan de ontwikkeling van jouw managementvaardigheden. Ik waardeer je openheid en betrokkenheid. Clara, dank voor de geweldige coördinatie van het publicatieproces. Marije, Roos en Jos als mijn teamgenoten ben ik jullie dankbaar voor onze samenwerking, constructief en tegelijkertijd gezellig. Laurens, Elizabeth en Edwin, dank voor jullie samenwerking, adviezen en feedback op een aantal van mijn ideeën en conceptteksten. Marlies, Nanda en Sjerhiel, jullie namen mogen hier ook niet ontbreken, dank voor jullie hulp en gezelligheid. Oud-collega's Maaike en Floortje, jullie collegialiteit en vriendschap hebben zeker ook bijgedragen tot dit proefschrift. Dank voor jullie ondersteuning, humor en relativeringsvermogen.

De SciSA junioren verdienen een speciale plek in dit dankwoord: Bei, Keelie, Pieter, Roos, Stefan, Thomas en Tjerk, jullie maakten het Rathenau tot een bijzondere werkplek, waar hard werd

gewerkt en ook hard werd gelachen. Beide waren nodig om te komen tot dit proefschrift. En onze JuSci-meetings niet te vergeten: het delen van onze *moments of happiness* en moments of *despair*. De meesten van jullie hebben inmiddels een andere werkplek gevonden en ik wens jullie allen heel veel succes, geluk en plezier! Stefan and Thomas, my 207 roomies, you are two very special men who taught me so many valuable things. Not only about you and your colorful lives, about modern technology, music, food and how to win a discussion, but especially about myself. I am very thankful for your friendship and support, your sense of humor, incredible openness and even greater stubbornness. I will not easily forget our fruit breaks, gym sessions, dinner dates and city trips and I hope many more will follow!

Dank ook aan Christine Teelken, Dick de Gilder en Peter Groenewegen, collega's van de Vrije Universiteit, vakgroep Organisatiewetenschappen, met wie ik vooral in de beginfase veel ideeën heb uitgewisseld en die me geholpen hebben bij het vormgeven van mijn onderzoek.

Mijn lieve familie, schoonfamilie en vrienden wil ik bedanken voor hun steun en gewenste afleiding door de jaren heen. Jullie hebben mij met beide benen op de grond weten te houden en me geholpen werk af en toe helemaal te vergeten. Uitleggen wat ik nu precies deed in Den Haag, werkte daarnaast ook erg relativerend. Mijn ouders wil ik graag bedanken voor wie ze zijn en wat ze mij hebben meegegeven. Hoe meer ik mij ertegen verzet, des te meer herken ik mezelf in jullie. Maar één ding is zeker, zonder jullie ongelooflijke koppigheid was dit proefschrift er niet geweest. Een persoon in het bijzonder wil ik hier nog benoemen en dat is oma Van Dijk. Waar ik het in mijn vrije tijd vaak heerlijk vond om even niet over mijn werk te praten, vroeg jij juist altijd naar mijn werk. En dan niet een algemene vraag zoals 'hoe is het op je werk', maar je wilde precies weten wat ik nu deed, hoe, waarom, met wie en vooral wanneer ik klaar zou zijn voor de verdediging. Zo trots als je altijd was, zo graag had ik het je gegund tijdens mijn promotie vooraan te kunnen stralen. Je had het vervoer en reisgezelschap al geregeld, helaas piepte je er net te vroeg tussenuit. Oma, dit proefschrift is voor jou.

Roel, jij gaf me de ruimte om volledig op te kunnen gaan in mijn werk en je gaf me op de juiste momenten een relativerende schop onder mijn kont om weer te beseffen dat er belangrijkere dingen zijn, zoals onze prachtige zonen Mees en Simon, jij en ik. Ik ben je ontzettend dankbaar voor je vertrouwen, geduld, toewijding, humor en liefde. Lief, je bent mijn alles.

**Who was Rathenau?**

The Rathenau Instituut is named after Professor G.W. Rathenau (1911-1989), who was successively professor of experimental physics at the University of Amsterdam, director of the Philips Physics Laboratory in Eindhoven, and a member of the Scientific Advisory Council on Government Policy. He achieved national fame as chairman of the commission formed in 1978 to investigate the societal implications of micro-electronics. One of the commission's recommendations was that there should be ongoing and systematic monitoring of the societal significance of all technological advances. Rathenau's activities led to the foundation of the Netherlands Organization for Technology Assessment (NOTA) in 1986. On 2 June 1994, this organization was renamed 'the Rathenau Instituut'.

The quality of higher education and research is strongly connected to the quality of the people working in the academic sector. For excellent science, excellent scientists are needed. The pool of competent scholars with academic career ambitions has been growing for the past decades. As public funding of universities has not matched this rise of (potential) staff, academics have become more dependent on competitive external project funding and individual funding. Publishing and grant proposal writing are two activities of major importance for especially early career researcher who aim to advance in academia. Funding organizations therefore play a crucial role in the development of academic careers, next to universities that increasingly focus on attracting and retaining academic top talent. The research questions of this study, 'What is academic talent and how is it selected?' aim to create a better understanding of the process of talent selection within academia, especially in the context of grant allocation.

Key results of this study address the criteria used in talent assessment and more specifically the weight assigned to publications; the social and competitive nature of grant allocation processes; the role of gender in talent selection and gender differences in academic performance; and factors supporting or impeding academic careers.

This study feeds current debates on scientific quality and the growing competition for funding and academic positions with empirical arguments. It reflects on the existing mechanisms of talent selection and ends with a discussion on the implications for higher education and science policy to uphold and stimulate academic talent.

ISBN 978-90-7736-457-4