RESPONSIBLE RESEARCH DATA MANAGEMENT AND THE PREVENTION OF SCIENTIFIC MISCONDUCT

Royal Netherlands Academy of Arts and Sciences
PO Box 19121, NL-1000 GC Amsterdam
T +31 (0)20 551 0700
F +31 (0)20 620 4941
knaw@knaw.nl
www.knaw.nl

pdf available on www.knaw.nl

Basic design: edenspiekermann, Amsterdam
Typesetting: Ellen Bouma, Alkmaar
Translation: Balance Amsterdam/Maastricht

Illustration cover: iStockphoto

ISBN 978-90-6984-656-9

The paper for this publication complies with the ∞ iso 9706 standard (1994) for permanent (durable) paper

RESPONSIBLE RESEARCH DATA MANAGEMENT AND THE PREVENTION OF SCIENTIFIC MISCONDUCT

Royal Netherlands Academy of Arts and Sciences
Advisory Report by the Committee on Scientific Research Data,
April 2013

FOREWORD

The Royal Netherlands Academy of Arts and Sciences (KNAW, "the Academy") has always been an enthusiastic advocate of open access to research data and research results. Maximum access to data supports pre-eminently scientific methods in which researchers check one another's findings and build critically on one another's work. In recent years, advances in information and communication technology (ICT) have been a major contributing factor in the free movement of data and results. It is against this background, but also in the light of recent cases of research fraud, that the Academy has undertaken to investigate how various disciplines actually deal with research data and to consider whether their practices are satisfactory. The Academy has entrusted this important task to an ad hoc advisory committee chaired by Prof. Kees Schuyt.

I heartily endorse the recommendations that the Committee makes in the present report. I would like to emphasise five of those recommendations in my own words:

- Allow the various scientific disciplines to decide for themselves the best way to deal with research data, but make free availability of that data the default option.
- The research community does not so much require additional rules of conduct but, above all, measures to "revitalise" the existing rules.
- When evaluating research, we should investigate how the research data is dealt with and make suggestions for revitalising the rules of conduct.
- We should consider the extent to which we can prevent scientific misconduct by studying best research practice.
- Finally: it is a privilege to be scientist. Let us work together to ensure that our profession remains enjoyable. As I have already indicated, we should do so not by agreeing on more rules but by concentrating on responsible research conduct.

The Academy owes a great debt of thanks to the Committee on Scientific Research Data. It is now up to the reader to draw conclusions. The report's recommendations are addressed to the relevant members of the research community, and I trust that they will understand what is required of them.

Prof. Hans Clevers
President of the Royal Netherlands Academy of Arts and Sciences

CONTENTS

FOREWORD 5

1. INTRODUCTION 9
1.1 Background to establishing the Academy's Committee on Scientific Research Data 9
1.2 Remit of the Committee 10
1.3 Interpretation of the Committee's tasks 10
1.4 Responsible research conduct 11
1.5 Integrity 12
1.6 Borderline cases and the critical zone of questionable practices 12
1.7 Contents of this Advisory Report 13

2. RESPONSIBLE RESEARCH DATA MANAGEMENT 16
2.1 Introduction: as a precaution 16
2.2 An analytical survey of research practices 17
2.3 Differences between and within scientific fields 19
A. Natural sciences (in the broad sense) 21
B. Medical and biomedical research 22
C. Mathematics, logic and philosophy, humanities 24
D. History 24
E. Agricultural sciences 25
F. Behavioural sciences and social sciences (I: economics, econometrics, and social geography) 26
G. Behavioural sciences and social sciences (II: including sociology, political science, psychology, and educational theory) 28
2.4 The research cycle: three assessment points and the role of the scientific forum 31
2.5 Data management practices 33
2.6 Corresponding features of good practices 39
2.7 The Committee's main recommendation 40
2.8 Recommendations 41

3. INTEGRITY IN SCIENTIFIC RESEARCH 47
3.1 Integrity as a guiding value in science 47
3.2 Scientific fraud as a warning 50
3.3 A spectrum of scientific misconduct 52
3.4 Is there a relationship between research data management and scientific misconduct? 55
3.5 Scientific misconduct and the system of scientific practice 56
3.6 Recommendations 60

4. PREPARATION FOR SCIENTIFIC RESEARCH IN UNDERGRADUATE AND POSTGRADUATE PROGRAMMES AND DURING RESEARCHER TRAINING 65
4.1 Introduction 65
4.2 Undergraduate and postgraduate programmes 66
4.3 Researcher training 67
4.4 Recommendations 68

5. SUMMARY 69

BIBLIOGRAPHY 77

APPENDICES
1. Resolution inaugurating the Committee on Scientific Research Data 79
2. Peer reviewers 81
3. Interviews and written replies 82

1. INTRODUCTION

1.1 Background to establishing the Academy's Committee on Scientific Research Data

This Advisory Report was produced by the executive officers of the Committee on Scientific Research Data, established by the Royal Netherlands Academy of Arts and Sciences. A thorough understanding of its contents requires readers to know the background to the Report. The Academy had two very different reasons for establishing the Committee.

The first reason was the Academy's wish to investigate how such trends as digitisation and the internationalisation of research might offer new ways of improving access to research data for other researchers (data sharing). The Academy organised a conference on 12 December 2011 to discuss ways of improving data management, storage, corroboration, and consistency and other issues connected with storage and management (sharing and control). The conference underlined the need for a survey of current data storage and data processing practices within the various scientific fields.

This intention on the part of the Academy overlapped with a second reason for establishing the Committee, namely a spectacular case of research fraud at Tilburg University (the "Diederik Stapel case"). That case caused tremendous uproar in the media and raised questions about the reliability of scientific research. Public confidence in science was shaken. Questions were raised in Parliament and journalists set out to determine whether this was an isolated case of a researcher inventing or himself filling in research data, or whether it might be the tip of an iceberg in Dutch research. The research world and the universities were expected to respond to the various questions raised: how had such a case been possible despite science's self-refining ability? The Stapel case led to a broader remit for the Committee on Scientific Research Data.

1.2 Remit of the Committee

In the light of the above matters, the Committee on Scientific Research Data began its work with a broader remit. It was also asked to carry out its work within a very limited period of time.

The Committee's main task was to produce a survey of data acquisition, storage and management, analysis, and monitoring practices within the various disciplines, and of how data can be archived for purposes of verification and perhaps for later research. The Committee was also asked to make recommendations conducive to responsible data management. The Resolution inaugurating the Committee was taken because "…discussion has arisen within the research community and society in general regarding how researchers deal with research data". The instructions issued to the Committee also refer to scientific integrity: "The task of the Committee is to draw up recommendations encouraging researchers in all disciplines to familiarise themselves with routines that promote scientific integrity in their data management practices" (Resolution inaugurating the Committee, Appendix 1). A third component of the Committee's remit was also specified: "the responsibilities of researchers and their employers (including universities and Academy research institutes) for disseminating and complying with standards of scientific integrity, including the teaching and supervision of young researchers".

1.3 Interpretation of the Committee's tasks

The Committee interprets this tripartite remit as follows:

1. to survey existing data management practices in various scientific fields;
2. to identify workplace routines that encourage researchers to act with integrity in their work;
3. to designate and allocate responsibility for communicating standards of scientific integrity to young researchers (in the context of teaching and supervision).

In the view of the Committee, responsible research data management and research integrity are closely related. If a researcher's conduct is irresponsible, there is a greater likelihood that his or her research will lack integrity. Responsible conduct and integrity are both inherent to good research practice and form the core of science. The authority of science and confidence in research results both depend on responsible research conduct and research integrity, as do mutual trust between researchers and the possibility of building on one another's work. The scientific ethos is governed by a readiness and willingness to report truthfully on the research process so that the results can be tested by the research community, by fellow researchers, and by critical recipients of the results. When confidence in research is severely tested or even betrayed – for example in cases of serious fraud – the researchers involved have almost always conducted themselves irresponsibly and without integrity by inventing data (apparently an uncommon occurrence) or by deliberately omitting, incorrectly representing, or stretching data (considered or asserted to be less uncommon).

The Committee wishes to point out, however, that responsible research conduct and scientific integrity are not identical. Responsible conduct is a matter of gradations, ranging from exemplary to good, mediocre, poor and sloppy research practice. Scientific integrity, however, is something that can be clearly defined. After a certain point, integrity is compromised and becomes research fraud.

The Committee's primary objective in the present Advisory Report is to cast light on responsible research data management practices. What can be considered responsible research data management in the various phases of research? It also wishes to investigate the relationship between responsible/irresponsible conduct and misconduct.

For the sake of clarity and as background to the Committee's advice, this introduction will first consider the two features of research practice already referred to, namely responsible conduct and integrity. In the view of the Committee, the two can be considered separate and independent categories of scientific behaviour; as such, they require separate analysis, discussion, and recommendations. The introduction also considers borderline cases and the critical zone of debatable practices; these subjects are discussed in greater detail in subsequent chapters.

1.4 Responsible research conduct

Anyone reviewing normal research practice in various scientific fields will be able to distinguish good practices – i. e. those that promote responsible research conduct – from less good or bad practices. Responsible research should in fact refer to the ability to justify conclusions on the basis of the data acquired and generated through research *and* the ability to account for that data; both features are an intrinsic component of responsible research practice. Accountability involves reporting on the research, for example in articles or papers, but the researcher may also be required to account for his or her work at a later date. Responsible data management therefore also means ensuring that research data remains available and, if necessary, is furnishing to other researchers for scrutiny and/or verification. If data is to be available for scrutiny, then proper archiving is necessary.

Research practices can vary from extremely good and reliable to poor and sloppy. The task of the Committee is to survey these practices in the various scientific fields, with comparisons between the different practices potentially serving to bring about any necessary improvements.

These practices can involve outstanding research in which researchers achieve creative breakthroughs and new insights; in other cases, those insights may not be based – or at least not yet – on responsible data logging or on precisely verifiable measurements. The requirement of responsible conduct may sometimes be at odds with scientific creativity. Researchers must be able to give their creativity and imagination free rein. In such cases, whether the new insights prove lasting depends on scientific discussion, verification, and scrutiny.

There may, however, also be cases of *poor* research: researchers may – in good faith – perform measurements incorrectly, something that becomes apparent when the data is checked or presented to the research community. The underlying reasoning may also be faulty, for example because it includes assertions that can be refuted. Poor research must be identified, opposed and improved as much as possible, things that a survey of research practices can help achieve. But although poor research cannot be equated with research *mis*conduct, it is always difficult to decide at which point irresponsible conduct enters the critical zone between vulnerable and doubtful or questionable practices.

1.5 Integrity

There is a yawning gap between research integrity – whether or not it is good research – and scientific misconduct: "If a scientist is suspected of falsifying or inventing evidence to promote his material interests or to corroborate a pet hypothesis, he is relegated to a kind of half-world separated from real life by a curtain of disbelief; for as with other human affairs, science can only proceed on a basis of confidence, so that scientists do not suspect each other of dishonesty or sharp practice, and believe each other unless there is very good reason to do otherwise" (Medawar 1984: 14–15). Researchers may cross the boundary into irrefutable dishonesty in serious cases of scientific misconduct, referred to internationally as "FFP": fabrication of data, falsification of data and data fraud , and plagiarism. All these types of misconduct undermine confidence in research, but not in the same way as poor research does; such misconduct is a more serious matter in which the honesty of scientific practice itself is at stake. Manifest dishonesty in obtaining research data and the falsification of data or presentation of false data lead to distrust of science precisely because the basis underlying the scientific ethos has been abandoned.

Strictly speaking, plagiarism might be defined as the dishonest reproduction of data rather than data falsification or fraud. Plagiarism, however, undermines the *system of rewards* in science and its venerable principle of giving "credit where credit is due". It destroys mutual confidence between researchers, although plagiarism does not necessarily mean that the recipients of the research results are actually presented with *false information*.

1.6 Borderline cases and the critical zone of questionable practices

Even if we accept that responsible research conduct and scientific integrity constitute separate categories, at what point does a researcher who mismanages research data contravene the standards that apply to scientific practice? There is a critical zone of "questionable research practices" in which research data management clearly leaves something to be desired, but in which it is generally not immediately clear whether the questionable practice is the result of sloppiness and a lack of corroboration (irresponsible conduct) or of a lack of integrity and dishonest intentions (misconduct). Such research need not necessarily be intended to mislead others but it can still call into question and undermine confidence in science. This critical zone of questionable research practices raises questions about precisely when and where the boundaries of scientific ethics are exceeded. Not every poorly executed study is an example of scientific misconduct. However, poorly designed and executed research produces a tremendous amount of "noise" in the dissemination of knowledge. This borders on scientific misconduct and in some circumstances – for example gross negligence or culpable irresponsible conduct – crosses that boundary. Such research is therefore unacceptable from an ethical point of view. In practice, however, these categories tend to blend into one another. Imposing boundaries creates borderline cases and creates a "grey zone" of behavioural types that can be assessed in different ways depending on the field concerned and the position of the researcher. Any gaps that arise in the monitoring of data collection or in the supervision of the research are attributable to the supervisors or monitoring bodies concerned, even though each individual researcher remains responsible for his or her own data.

There are two distinct cases that fall into this critical zone: (1) the researcher is honest but does not maintain good research practices, and (2) the research is so irresponsibly conducted that the researcher's integrity is at risk. In reality, the level of integrity and the researcher's intentions are often difficult to ascertain. One thing is clear, however: questionable practices need to be prevented and can in any case be improved. If they are the result of certain practices considered normal within the field or at the research institute concerned, then they need to be identified, discussed, and remedied as soon as possible in order to maintain (or improve) public confidence in science.

1.7 Contents of this Advisory Report

To summarise, this Advisory Report addresses two different aspects of research data management, both of which can be detrimental to public confidence in science: the extent to which research conduct is responsible (i.e. practices ranging from good to bad) and scientific integrity (the notorious cases of "FFP"). Each section also considers a grey zone of "questionable research practices"; here, it is not immediately apparent whether a particular practice is bad, incorrect, or dishonest. In all cases, the report indicates whether improvements are possible and whether they are necessary.

The Committee has drawn up a provisional outline survey of research data management practices. Unsurprisingly, these practices turn out to differ considerably from one scientific field to another. Section 2 reports on this and makes a number of recommendations. Section 3 considers the integrity of research and researchers and serious forms of ,scientific misconduct, and recommends various research practices for combating such misconduct and for identifying it as quickly as possible. Section 4 concerns itself with the third task of the Committee: to communicate standards for responsible research conduct and scientific integrity in education and in research practice. Section 5 offers a summary of the report.

In accordance with the Academy's quality standards for advisory reports, this report has been subjected to peer review. The names of the peer reviewers are given in Appendix 2.

Finally, the Committee wishes to point out the limited scope of the present Advisory Report. To begin with, it offers *advice on policy* and is therefore not a scientific report. It was not possible for the Committee to undertake a large-scale scientific study of research data management. The Committee therefore decided to conduct an oral and written survey of the personal views of researchers – ranging from junior to senior level – regarding research practices in a number of scientific fields. Although it has attempted to include a wide variety of disciplines and levels of seniority, the Committee has not aimed to achieve the kind of representativeness that would be required in an actual scientific study. The Committee does recommend, however, that a randomised study should be carried out in future to produce a representative picture of the various research practices (see the recommendations in Sections 2.8.2 and 3.6 of this report).

Second, the Committee is offering a mere "snapshot"; it is aware that various trends within the world of science make constant consideration of the quality of research necessary. These trends include the ongoing digitisation of research data; the new possibilities made available by the Internet, including new data acquisition options; and the changing relationship between pure and applied research and between unsponsored and sponsored research.

Third, the Committee's remit is explicitly limited to *publicly funded and co-financed research* carried out by researchers associated with public organisations. The percentage of research carried out in the Netherlands by businesses, in company laboratories, workplaces, and innovation centres is not insignificant. Increasingly, university-based research and commercial research are pairing up, with a growing proportion of research being commissioned or sponsored by private enterprise (KNAW 2005). Similar questions can be posed concerning the management, storage, corroboration, and control of research data obtained through private research and in research commissioned by the private sector. The Committee also realises that the needs of public and private (or privately financed) research do not differ significantly when it comes to responsible research conduct and scientific integrity. After all, good research means the same thing in both fields. Nevertheless, research financed by third parties and contract research does raise specific questions that justify more extensive consideration, for example patent applications, contractual conditions regarding the use and ownership of data, publication rights and publication periods. Where these separate problems raised by contract research are concerned, the Committee refers to the study carried out by another Academy committee: *Wetenschap op bestelling* [Science to Order] (KNAW 2005).

Note: This report makes use of the masculine form; this should be taken to refer to both male and female researchers. Similarly, "science" and "scientific" should be taken as referring broadly to both science and scholarship, i.e. the humanities and the social sciences in addition to the natural sciences.

2. RESPONSIBLE RESEARCH DATA MANAGEMENT

2.1 Introduction: as a precaution

Should the research community be worried about the overall quality of scientific research, in particular the manner in which researchers manage their data? The Committee set out to consider that question because regular reports of questionable research practices have put confidence in science at risk. Public distrust of science appears to be growing, fed by a series of incidents that have received considerable play in the media. Nevertheless, it is unclear just how significant those incidents actually are. "Is this the tip of the iceberg?", journalists ask, without immediately being able to answer that question. Politicians have questioned the State Secretary for Education, Culture and Science, while the Association of Universities in the Netherlands (VSNU) and the Royal Netherlands Academy of Arts and Sciences – as the organisations responsible – have felt called upon to respond and have decided to take such warnings seriously. Both in the world of science and the public are worried, in other words – without our being able to say precisely what the object of that worry is.

The position of the research community's itself on these matters is more reassuring, however. Broadly speaking, that position – generally described by the researchers themselves – is as follows:

> *Research results are so stringently monitored in most scientific fields that poor and sloppy research is filtered out. The growing pressure on research results and on the quality of scientific publications means that the internal, self-refining ability of science is more than sufficient to guarantee the quality of research. This stringent testing means that scientific research can still be trusted fully.*

But is that self-image in fact correct? Do researchers not have a tendency to project an image of science to the world outside that is finer, better, and stronger than is justifiable? Without a full-scale scientific study, it is extremely difficult to appraise either the alarming or the reassuring reports. Such a study would need to cover not only questionable or dubious research practices but also standard, unquestioned practices, and even exceptionally good practices. Only then can the ratio of less good to standard research practices become clear (see Sections 2.8.2 and 3.6). But leaving aside the need for such a study, the Committee wishes to note another viewpoint relevant to the confusion concerning the state of scientific research, namely the precautionary principle. That principle suggests that the world of science must combat every realistic possibility that confidence in science will be undermined, and it must do so as rigorously as possible and at the earliest opportunity. Even if there are no direct signs pointing to fraudulent research practices, there is every reason to investigate whether the practice of scientific research is such as to remove any doubt as to its quality and integrity. The precautionary principle acts as a kind of tall, solid dyke that removes the risk of flooding and reassures the inhabitants of the delta. It is for precautionary reasons that the scientific world must ensure continuing confidence in science and scientific research by asking itself whether researchers are in fact complying with all stringent requirements. In other words: be fully prepared so that the country's inhabitants know that they really *are* safe behind the "dyke" and do not need to worry about flooding. Given the doubts that have arisen among the public, an appraisal of current research practice – no matter how provisional – is both necessary and justified. Is the scientific "dyke" being monitored as thoroughly and effectively as it could be?

2.2 An analytical survey of research practices

In order to identify everyday research practices in the various scientific fields and recognise the points at which risks are involved, the Committee held two hearings at which it conducted fifteen interviews with representatives of various different fields (PhD candidates, postdocs, research coordinators, professors). These "hearings" were also continued in written form by sending a list of seven questions (to be answered in writing) to a group of researchers representing a larger number of fields. This group also included PhD candidates, postdocs, research directors, professors, and representatives of NWO's Divisions and the Academy's Advisory Councils). As pointed out in Section 1, this survey does not aim to be representative the way a proper large-scale study would attempt to be. The number of people approached by the Committee is much too small to represent all the views within the various scientific fields. The Committee's therefore did not attempt to quantify those views but merely to gain an overall picture. Seventy-nine responses were returned with extensive answers and opinions on research practice in the respondents' fields; they covered eleven major scientific fields and the separate disciplines within them (see Appendix 3 for a list of the fields involved and the researchers' job titles).

The following questions were posed both at the hearings and in the written interviews:

1. Are there questions or concerns in your field about how researchers manage their data (collection, processing, statistical analysis, verification)?
2. Is data management monitored, and what does that monitoring involve?
3. What happens to the research data after the research concludes (archiving, possibility of replication)?
4. Where in the research cycle – i.e. from the data collection phase right through to publication – is there the greatest risk of something going wrong? When *do* things go wrong in your opinion?
5. Are there sufficient monitoring mechanisms in place for data management? How should this be arranged: separately for individual researchers (or groups of researchers), or on an institute-wide basis?
6. How can open access to data be ensured in connection with data sharing? Is it in fact possible and/or desirable for data to be openly accessible?
7. How do we "keep it fun"? In other words, how can we improve the quality of research data management without introducing more bureaucratic rules? How can we avoid bureaucracy in research and keep research from becoming bureaucratic?

The answers to these questions – which are in themselves simple ones – were often surprising and generated various suggestions for improvements in research practice. Researchers were very willing to answer the questions, a sign that they are themselves concerned about the issue of confidence in science. As could be expected, the responses differed from one field to another. However, even within fields, opinions varied on various aspects of research data management. Sections 2.3 and 2.5 report on the survey.

This survey of research practices enabled the Committee to formulate four conclusions regarding the problems involved in research data management. The conclusions are explained in Sections 2.3 to 2.6. The survey also led to recommendations for improving data management practices. The Committee's main recommendation (Section 2.7) is based on those conclusions. The recommendations from the present section are presented in Section 2.8.

The four conclusions are as follows:

1. The type of research differs so much both between and within the various fields, and the research practices are so diverse, that it is pointless to make *general statements regarding the quality of research data management,* nor is it possible to determine the extent to which researchers observe good or best research practices.

2. One phase of the scientific research cycle that is *relatively free of external monitoring* and that offers huge scope for creativity is the primary research process (i.e. the process of acquiring, storing, organising, processing, and accounting for data),

in other words after the *start of the study* and before the *peer review.* Depending on the field involved, this is a high-risk phase in terms of research data management. Those risks may be due to shortcomings in monitoring, although the existence and nature of these shortcomings vary from one field to another. A similar gap in monitoring can be found in some fields during the phase when research data is archived after the end of the study. The absence of proper archiving makes it difficult to corroborate the data at a later date. Although the usual round of criticism and discussion within the scientific forum lowers the risk, weaknesses in the peer review system can nevertheless be identified in a number of fields.

3. With so many differences between and within scientific fields, it would be best to identify the risks, shortcomings in monitoring, and verification improvement options for the relatively exploratory initial phase of the research cycle *in each separate discipline*. If monitoring "after the fact" – i.e. peer review, scientific forum – is troublesome and not watertight, the obvious solution is to look more closely at the "before the fact" phase. Gaps in monitoring can also best be *identified* for each separate discipline. The various disciplines can learn from the good practices of other disciplines and – if necessary and possible – take over various monitoring mechanisms from those disciplines (for example keeping diaries and logbooks or lab journals, accounting for data, carrying out research within teams, and ensuring peer pressure prior to the peer review phase).

4. Good data management practices that are already established within a number of scientific fields can, where necessary and applicable, be introduced within other fields. Examples include arrangements and protocols in international research; established procedures for collecting, managing, storing, and using research data in large-scale research; and the inclusion of external monitoring mechanisms in research that is frequently carried out by individual researchers working in isolation, owing to the nature of the subject.

The following subsections explain and elaborate on these four conclusions.

2.3 Differences between and within scientific fields

Dutch researchers carry out a huge amount of research, some of it being of world quality. That research varies enormously in scale and is extremely diverse, making it challenging to obtain an accurate picture. The country has 75 accredited research schools. These vary in size, but it is not unusual for them to employ some 150 to 250 researchers; half of these are generally PhD candidates. University medical centres sometimes have more than 400 PhD candidates (doctors undergoing training and studying for their PhD). In 2012, the Dutch PhD Candidate Network [*Promovendi Netwerk Nederland, PNN*] had some 9000 PhD candidates as members, and there are also a large number of external PhD candidates engaged in research (most of them part time). Much of the scientific research carried out in the Netherlands is therefore conducted by PhD candidates. A considerable amount of research is also carried out by postdoctoral students, researchers funded by the Netherlands Organisation for Scientific Research (NWO), university lecturers ("UDs"), senior university lecturers ("UHDs"), and professors. The total research capacity at Dutch universities in 2012 has been estimated at approximately 17,000 full-time jobs ("FTEs") (Chiong Meza 2012).

All these researchers manage data in some way or another, although the kind of data differs enormously, ranging from astronomical measurements, stem cell research, computer simulations of buildings, and annual trading stock figures right through to the analysis and interpretation of a single poem. Even when large-scale data collection forms the core of the discipline, there can be enormous variety. In astronomy, large observatories and space agencies have long had standard procedures for storing and managing data (which is generally obtained in the context of international collaboration), managing accessibility, controlling who accesses data and when they do so before it is ultimately released for general use. These procedures differ considerably from those that apply in physics (for example particle physics), which often involves collaboration between consortiums made up of hundreds of research institutes, where the data generated is often only processed and reported on within and by the consortium. It is no longer unusual for an article to have more than a hundred authors in this field. The nature of biological research has changed due to the many computer applications now involved. Besides systematic observation and recording, computer simulation is also possible now, a development that has changed data management in this field and the extent to which data and analysis can be monitored.

There is nothing new in observing that the traditional scientific fields – the humanities, the physical sciences, and the social sciences – differ considerably. It is worth noting, however, that the different scientific fields differ so much from one another and internally that any recommendations regarding data management must allow for the specific features of each discipline. Not only do disciplines differ in terms of the *object* of research but also in terms of the *type* of research performed, the *type of data* collected or acquired, and the extent to which data management *agreements and protocols* have already been introduced as standard procedure (sometimes a considerable time ago). The responses to the survey confirm this variety. The present section discusses the consequences of these differences in light of the Committee's remit, namely to recommend refinements or, where necessary, improvements in research data management methods. A brief sketch of the main features of a number of scientific fields will be followed by the Committee's conclusion that the variety between and within the various fields is too great to make general statements or provide general recommendations that apply to all these fields.

A. Natural sciences (in the broad sense)

Representatives of the natural sciences responded to the question of whether they worried about research data management – i.e. data collection, processing, management, and analysis– with a resolute "no". Indeed, the question came as a surprise to them, as expressed by a leading representative of this field in his written reply:

> *There are no questions or worries in a general sense. It is the task of every supervisor of PhD candidates, postdocs, and other junior researchers to instruct them in the art of correctly managing and processing their measurement data. The self-correcting nature of the publication cycle – with publications being subject to strict peer review and ultimately approved (let us hope) – acts as a guarantee for the rest. As a peer reviewer, I have myself prevented the publication of results on a number of occasions when their analysis was substandard or when the statistical margin of error was too wide to draw valid conclusions (in other words, the authors' assertions were not justified by the quality of the experimental data or the margin of error).*

In this reply, responsible data management is "enforced" by stringent reproduction and verification practices. Another physicist refers to this kind of internal monitoring:

> *It should be said up front that fraud can never be excluded completely, as the Schön case shows. In virtually all cases, research data is collected by teams – large or small – of researchers, thus guaranteeing a certain level of internal monitoring. Analysis is often carried out by individual researchers within the team but the results of that analysis are in most cases subjected to critical discussion within the team and by larger bodies of researchers, for example research groups and departments. These discussions often throw up new research questions, with additional tests being carried out.*

In practice, the types of data and databases vary enormously, but every scientific field has developed a set of practices that comply with that field's requirements. The research group and the supervisors act as a fixed point of reference in that context, not so much as "sentries" but as an obvious point of contact and as an "extra pair of eyes" for the researcher. This applies in particular to physics and astronomy, chemistry, and in general to research in fields involving natural phenomena – earth sciences, biology, technical sciences – in which a long research tradition is accompanied by stringent external monitoring by the scientific forum. One might say that nature "strikes back" when a measurement has not been carried out properly and other researchers are unable to replicate the result. In other words, the later phases of the research cycle, combined with the nature of the discipline, act as a powerful restraint on poor performance and unverifiable data: the more research results are monitored later – including by means of the usual replication – the more the researchers will check them beforehand. The system functions extremely well, something demonstrated by the existing protocols on data submission and preparation.

However, the critical tradition in the exact sciences does not necessarily apply when it comes to monitoring *data management practices* in the same field. A common response to the interview question concerning such monitoring was that the individual researcher managed the data on his own computer and that there was no systematic data management policy. When researchers work with others within a larger research group, however, the supervisors (including PhD supervisors) are often responsible for monitoring. In these favourable circumstances, all the data – together with the associated lab journals or logbooks – is retained for a lengthy period of time, meaning that subsequent checks are always possible. Public availability of data and the possibilities for data sharing varied according to the nature of the field concerned. In certain fields, such public availability is properly arranged and there are fixed agreements, but in other fields researchers are more hesitant about sharing data with "outsiders".

In a few cases (physics, biology), respondents were concerned that researchers did not test out their ideas and hypotheses sufficiently and that – partly owing to pressure to publish quickly – they did not conduct enough experiments to be "entirely certain" of their results. Nevertheless, the respondents expected that incorrect results would soon come to light in such cases. Any doubts would be swiftly removed.

B. Medical and biomedical research

A different pattern has emerged in medical and biomedical research. First of all, an important distinction can be made in this field between laboratory research into mechanistic explanations of life processes and research on actual people. Research on people consists of both population studies (epidemiological, genetic) and clinical research on patients. Clinical research on patients may also be epidemiological or genetic, but it includes randomised controlled trials (RCTs) and clinical research (often on a smaller scale), for example in the context of PhD research within clinical departments.

All these categories feature specific research practices and databases. Research skills such as statistical analysis of populations or groups of clinical patients differ from the skills involved in fundamental cell biology research on stem cells, for example.

Laboratory monitoring differs from monitoring of large data collections. In the lab, checking involves not only regular observation and repetition of tests but also keeping a detailed lab journal of results that also explains why changes have been made to the test design. Clinical and population databases vary hugely. In some cases, monitoring is strict: checks are carried out whenever data is entered or altered, there are "site visits", and data is recorded over many years (in the case of RCTs, the research involved concerns the registration of medication). Data is also almost always managed centrally in the case of large-scale clinical, genetic, or epidemiological research (data managers updating the "parent files" and files derived from them). In the case of clinical research on a smaller scale – which makes up a large proportion of the research total – it is generally the PhD candidate himself who stores the data on his own PC, within small groups, and without much supervision by specialists in data collection and analysis.

One special characteristic of biomedical research in general is that it focuses on achieving "breakthroughs" that lead to more or better cures. There is an almost mandatory research imperative to prevent illness and death by constantly coming up with new findings and improved therapies (Callahan, 2003). The investment involved – both emotional and financial – is considerable, and that goes for both laboratory and clinical research. Controversies can escalate enormously – even ending up in court – and access to data can be restricted.

Clinical research involves specific aims and surroundings and specific ambitions on the part of individual researchers. A considerable amount of clinical research at medical centres is performed by PhD candidates who subsequently apply for training as a GP or specialist. Research is also carried out by specialists undergoing training with a view to obtaining a doctorate alongside their demanding clinical work with patients and their own specialist training. Most of these do not intend "going into research". Their clinical research is often supervised almost entirely by clinicians who have other, major responsibilities. These supervisors – who are often heads of department – are responsible for patient care, for training medical specialists, and for educating medical students (teaching and clinical placements). These problems have already been noted (Altman, 1994, 308:283) but have seemingly persisted.

All this ultimately involves a large volume of research that is widely distributed within the university medical centres ("UMCs"). Such multiple activities and fragmentation mean that research data management requires an extra level of attention. Although the clinical research question must continue to function as the key source of creativity, additional (statistical) research skills and support are at least as important here as scientific ingenuity. Consideration should also be given to time management, so that sufficient time is available to perform or supervise research. The current situation – which involves huge volumes of all kinds of clinical research – probably leads primarily to "noise" and not directly to fraud. In biomedical research, data management is generally monitored by conducting *audits* (the quality system applied by the Netherlands Institute for Accreditation in Healthcare or NIAZ) of the institution or research group concerned. This is rapidly developing field, with increasing use being made – or potentially being made – of data management software. Because medical research often takes place within the context of large organisations, those involved are acutely aware that systematic monitoring to ensure professional research data management can result in considerable extra costs.

Two physicians responding to the Committee's questions offer contrasting views, one confirming and one denying this description of biomedical research. The first says:

> *I believe that the research cycle consists of a large number of phases that are all equally weak. The greatest risk is that various aspects do not operate optimally at all kinds of levels (data storage, data processing, and analysis), meaning that the end*

product risks being mediocre at best. Otherwise, my impression is that most errors are "unintentional". In other words, deliberately false use of data is only sporadic.

By contrast, the other says:

Regular audits in the context of the NIAZ quality system, with clinical research also being involved. How this should be done is currently being investigated. A hospital-wide data management system is also being made available for inputting and managing research data and researchers are being trained. All of this is combined with researcher awareness-raising by means of BROK courses.[1] A form of monitoring will be introduced within the foreseeable future.

C. Mathematics, logic and philosophy, humanities

Respondents in some disciplines confidently denied there being worries or problems related to data management, simply because the "data" involved consists exclusively of publicly available texts (philosophy, theology, law, literary studies) or of analyses and reasoning whose validity can be tested or "recalculated" by other specialists in the discipline (mathematics, logic, analytical philosophy). In fact, the field concerned consists precisely of arguments put forward by specialists in publicly available, scholarly articles or discussions that can be checked by other specialists in the discipline.

Text analysis – for example in the context of literary studies, philosophy, and theology – or interpretation of legislation and case law – as in the context of legal studies – does, of course, throw up problems and discussions of interpretation but there is little to conceal and little in the way of data that is collected and managed solely by the relevant researcher. The public nature of a text, for example a poem by Shakespeare, or of a piece of philosophical reasoning, therefore acts as a restraint on irresponsible research. If a researcher acts irresponsibly, it generally means that he has been excessively selective in quoting texts that support his own interpretation or theory, something that will immediately be noticed and criticised in the course of discussion within the scholarly forum. The quality of the scholarship is made clear by public recognition of the reasoning that has been displayed and checked (although there were troubling cases of plagiarism recently in the field of legal studies in Germany). Research data management would therefore seem to be typical of empirically-driven disciplines but is only relevant in the humanities when they make use of such data taken from other scientific fields.

D. History

History and some of the humanities involve dealing responsibly with publicly available texts. Historical research consists of examining sources that are basically in the public

1 BROK: Basic Course on Regulation and Organisation of Clinical Research

domain. Often, however, only a handful of researchers actually delve into and consult these sources. Someone who has carried out years of archival research – for example on the actions of the Stasi in the former East Germany – can only be checked up on by another researcher who has also spent years consulting those same sources. Such verification is mainly undertaken for reasons of principle, and it is possible because all the sources are publicly available and accessible. The sources themselves may contain inaccurate information, but it is the task of the historian to identify and if necessary correct those inaccuracies. The quality of historical research depends on the researcher's conduct: has he conscientiously made those sources accessible for other researchers or presented them in such a way that others can assess the plausibility and originality of his research output? The responses historians have given to questions regarding data management monitoring reveal an interesting paradox. On the one hand, they are emphatic that data fraud is not possible and that it does not occur; on the other, they say that there is little monitoring of data management because they only make use of publicly available data. This basically constitutes a risk.

> *Historians don't collect data; they study existing data that can be found in archives or that is managed by documentation organisations. Researchers can make errors in interpreting data; they can be careless in their archive documentation; and there can even be errors in the archives themselves (in demographic research you can sometimes find dates that are off the mark by 130 years because of a clerical error). The following points clarify what is customary in terms of collecting, processing, statistically analysing, and verifying research data.*
>
> *Research in economic and social history involves not just published data – reports, published statistics, secondary literature – but also a large amount of unpublished archival material. There are few concerns about the reliability of the archival research carried out by our researchers. First of all, annotated research based on sources can always be checked. The researcher would be shown up if he were to present bogus archival documents. Second, historical research involves analysis that results in an interpretation: it is always possible to discuss how the source material – which is often scarce and incomplete – should be interpreted. Even the most scrupulous analysis can lead to conclusions that subsequent additional research later shows to be incorrect. So it is highly unlikely that someone would fake research data (and therefore risk being outed as a fraud). Factual data is indeed extremely important in research but it is not the only "imperative" determining the outcome; the way the researcher brings in the context and historiography is just as important.*

E. Agricultural sciences

The picture is again entirely different in the agricultural sciences. Respondents were not concerned. They have a great deal of practical data management experience. Data is collected under the auspices of established research institutes and then made available to researchers. Much of it is either publicly available or becomes so after publication. Data collection, storage, and processing are subject to joint supervision.

> *I am not aware of any concerns about research data management in the field (production, ecology, and agricultural sciences). The Agricultural Economics Research Institute (LEI) (part of Wageningen University and Research Centre or WUR) publishes an annual Agricultural Economics Report on the situation that contains a large quantity of strategic data. That information is freely available. Alterra ["the research institute for the green living environment"] also collects a lot of ecological data and observations of phenomena, and in general this is also freely available. The reliability of the data is subject to close scrutiny. The procedures are documented in protocols that have been developed over many years and that are utilised by researchers and those who make use of the data. These are standard research protocols, however. After the end of each study, the basic data remains available at least until publication has taken place in international journals.*

F. Behavioural sciences and social sciences (I: economics, econometrics, and social geography)

Confidence in scientific research is apparent from the responses provided in two different fields, namely econometrics and social geography, both of which work mainly with public and publicly accessible data. The economist responded that there were few data management problems in economics because modern economic and econometric research takes all its data from existing, publicly accessible sources (annual figures, trade flows, consumer behaviour, public finances), meaning that every researcher can be stringently monitored. Competition is so tough in this field that carelessness or negligence in data processing is immediately punished. In other words, *the nature of the data* ensures meticulous management, assisted by mathematical analysis that reduces the likelihood of multiple interpretations. The fact that extremely diverging theories subsequently arise within the scientific forum regarding such economic phenomena as the causes of economic crises, unemployment, or inflation does not make the data less precise or its accuracy less plausible:

> *I work primarily in econometrics, so I will respond with that discipline in mind. ...*[2]*A lot of data is public (stock exchanges, national accounts) and anyone can check the results. If the data is not in the public domain, then a lot of journals ask for it to be released so that it can be provided on the relevant websites. Young researchers often request existing data files so that they can subject them to their own analysis. It is quite normal for such data to be provided, but only AFTER publication in a*

2 Where necessary, the Committee has anonymised the response by replacing portions of texts with an ellipsis.

professional journal. Researchers sometimes upload data to their website. More often, data is updated – for example day-to-day share prices or figures provided by Statistics Netherlands about consumer spending – and it's customary for researchers to make use of the most recent data.

There has to be a system of self-regulation within the field. If I use a prediction model for GDP, for example, and claim great success and someone replicates my study and it's not successful, then I'll look pretty foolish. Nobody wants that. People are very sceptical about elegant results.

Finally, econometrists are very good at statistics. Most of the time, they can see at a glance whether there is something wrong with the results, i.e. whether they look too elegant or too neat. I think that a course in statistics should form part of every scientific curriculum, but that's just my personal view.

Even in economics, however, not all data is – or even can be – made public, for example company results or factory secrets. The growing interest in behavioural economics has led to an increasing number of experimental studies, carried out in partnership with psychologists, while questionnaire-based research – for example on consumer preferences, budgeting and spending patterns, and job search behaviour – has long been a customary method in economics. Methodological variety therefore applies to the whole of the discipline of economics:

At our research institute, we make use of a very wide range of research methods: experimental research, longitudinal case studies, surveys, quantitative modelling, etc. The points to consider and possible ways of improving data management differ greatly from one method to another. Validation is largely ex post, while writing, presenting, reviewing, and publishing the research. Our institute has five research programmes and organises a large number of related seminars and symposiums where PhD candidates and faculty present their research results.

In almost all cases, PhD candidates are supervised by at least two researchers, and they also work within one of the five research programmes.

There is also an increasing tendency to check one another's work at the "front end", in other words during data collection and processing. That's because our researchers are collaborating more and more within teams (including international teams) and because they utilise "central" research facilities that we provide as a research institute (for example the… and electronic survey tools).

Data collection is organised in a similar manner in the discipline of social geography, with public verification being regulated effectively:

Most of the research is quantitative and statistical. The data we use is similar to that used in explanatory sociology and micro-economics, but we tend to combine it more with "geo-referenced" data so that we can include the effects on behaviour of the spatial context and estimate the collective outcomes of behaviour on spatial organisation. Our data resembles the types of data utilised by ecologists/systems biologists, except that the "agents" within the models are often people (in households) or businesses (in companies).

In the quantitative line, we make a lot of use of registers and large secondary databases. We also use digital mapping data a lot. In addition, there is an increasing level of harmonisation at European level (and at global level, although to a far lesser extent). Data collection is therefore largely in the hands of professional organisations with their own quality assurance systems for data collection and verification, and the metadata is generally in good order.

G. Behavioural sciences and social sciences (II: including sociology, political science, psychology, and educational theory)

The pattern of research practices in the broad field covered by the other social sciences (behavioural sciences and social sciences II) is extremely varied. There are signs of a move towards systematic research validation, but not everywhere. As in some other fields, the picture is very varied. Researchers use various different data collection methods, for example behavioural observation and experimentation in educational theory and psychology; qualitative and quantitative questionnaire-based research in social psychology and sociology; and public socio-economic and socio-cultural data and participatory observation of population groups in anthropology. In the latter case, an individualist research culture makes it difficult to systematically check data or research in the field. The methods of statistical analysis utilised in these disciplines also vary enormously.

In the field of educational theory, data management in the context of long-term cohort studies is properly organised:

We have data collected years ago on videotape which is now also going to be digitised. In the case of a longitudinal study lasting more than 20 years, the raw data files might still reveal something new. We also have central computer drives on which most of the data files are archived even years after the PhD candidate who did the study has left us, especially because we sometimes unexpectedly decide to carry out a follow-up. In our field, we store data for at least five years after publication.

Sociology has a long tradition of storing and managing large databases – in the Steinmetz archive, the predecessor to DANS (Data Archiving and Networked Services) – and of scrupulous evaluation of data collections. Large-scale international social research programmes utilise protocols and have mandatory arrangements regarding data collection, management, and monitoring. Sociology and political science also have a long tradition of large, shared databases, for example the long-term Dutch Parliamentary Electoral Studies (NKO), and the National Kinship Panel Study (NKPS). The data in these databases is available to anyone, without any embargo or other conditions applying. Enormous care is exercised when generating the research data and the collaboration involved means that data management, data storage, statistical analysis, and other uses of the data are closely monitored. In addition to the large-scale collaborative research, however, individual researchers also conduct many smaller studies, and these are the object of certain concerns:

> *Surveys are an important research strategy in political science. Unfortunately, the on-response rates are rising and this worries researchers because it is detrimental to the random sample, the size of the N, etc. Where integrity is concerned, the risk is that those carrying out surveys will fill in questionnaires themselves so that they can achieve their targets.*
>
> *In the case of large-scale questionnaire-based studies such as the NKO, there are random checks to see whether the interviews have in fact been conducted (follow-up calls, mystery respondents). In the case of small-scale studies, for example the Parliament Study [Parlementsonderzoek], the interviews are also recorded.*
>
> *The biggest risk is when each phase in the research cycle is in the hands of just a single person, i.e. the individual conducting the survey, the coder, or the researcher. Every deviation from the protocol – for example a researcher filling in a questionnaire himself; individual interpretation of the coding instructions; entering weighting factors; combining response categories, etc. – can be detrimental to the research. I myself am most concerned about qualitative research carried out by individual researchers (case studies, participatory observation).*

In psychology, data management is organised in a wide variety of ways. For one thing, the various divisions of psychology – experimental psychology, neuropsychology, developmental psychology, clinical psychology, social and organisational psychology – utilise many different types of research data, and they also use a huge variety of methods to collect that data, ranging from scanning and laboratory experiments to "pen-and-paper" questionnaire-based studies and large-scale surveys. Researchers carrying out laboratory experiments are supervised and monitored by research coordinators, who use a wide range of different methods. In some cases, researchers submitting studies for publication in an international journals must retain all the relevant essential research data in a separate dataset (i.e. a folder) or even submit that data along with the article so that it can be verified and accounted for. Other institutes, however, allow a single researcher or a handful of researchers to perform experimental studies in psychology laboratories with only occasional supervision. There is no general agreement on how to properly assess research data management practices within the same field. Some respondents referred to the limited monitoring in the psychology laboratory setting; others do not see problems because all the research is properly documented and all the data is accounted for in folders ("The Diederik Stapel case is just a one-off incident that could have happened anywhere; it's not specific to this discipline."). A third researcher is mainly worried about inadequate compliance with existing rules, both those of the international professional body representing psychologists and those imposed by the editors of academic journals. One of the respondents in this discipline noted that university programmes scarcely devote any time at all to teaching students research data management skills. Many psychology researchers would seem to confuse archiving with storing data on their own computer (i.e. the one they happen to be using at the moment). It seems that a culture of proper data management exists only in the case of large-scale longitudinal research.

> *Proper documentation of the research data prior to analysis is essential. In psychology, and particularly in experimental studies in the laboratory, it is quite normal for a single individual to be entirely responsible for data collection and processing. This opens the door not only to improper conduct but also to distortions, for example due to errors and an overly positive presentation of results. Moreover, failing to properly document data when it is still "fresh" often means that it is impossible to replicate one's analyses at a later stage (when a lot of implicit knowledge has faded away). Poor documentation limits the extent to which research data can be shared with other researchers after publication. It would already be a major improvement to make it standard procedure for researchers to share their data with one or more colleagues or co-authors for verification of their analyses (this is referred to as the "co-pilot model"). This would require thorough documentation of the data early on and it would ensure that the data can be recovered from a number of different locations. It would also prevent errors and reduce the likelihood of fraud.*

Some fields of the social sciences appear capable of learning from trends and developments in other fields – lab journals, researchers checking one another, protocols – and from one another (data archiving in sociology as a model for those cases in psychology in which such archiving is still absent).

FIRST CONCLUSION

The Committee's first conclusion is obvious but nevertheless important:

The type of research varies so much both between and within the different fields, and the research practices are so diverse that general statements regarding the quality of research data management are pointless and would say nothing about responsible research conduct per se.

General statements about the situation in "science" are untenable. Possible problems in research data management need to be investigated and discussed within the specific field concerned. Proposed improvements should also allow for the specific features of that field. A *general discussion* of research data in science can only be useful in tracking down weaknesses in the research cycle in a given field and to serve as an instructive example.

2.4 The research cycle: three assessment points and the role of the scientific forum

Given the major differences in research practice, we must seriously question whether certain general features of scientific research can nevertheless be identified. There is tension between collaboration and competition in every field; on the one hand, researchers want to be the first to present new knowledge and insights; on the other, they adhere to the venerable principle that the free exchange of knowledge and data is one of the best ways of ensuring scientific progress. Every researcher "stands on the shoulders of giants" and can see more if he collaborates with others and allows others to collaborate with him. Moreover, most fields of science have developed standard assessment procedures – for example for evaluating and accepting grant applications and assessing scientific publications by means of peer review – so that quality assurance mechanisms operate effectively in these fields and can teach valuable lessons. Finally, viewed abstractly, most scientific disciplines proceed according to what is basically the same "research cycle", meaning that we can attempt to identify the strong and weak phases of the cycle for each separate field, as well as the high-risk points. Comparing them reveals a number of common features or analogies, for example concerning the scale of the research: large-scale, internationally organised data acquisition, management, and archiving as opposed to individual research involving data that is mainly managed, analysed, and stored by the researcher himself.

The Committee has chosen to emphasise research data management within the primary research process by looking at the key assessment points in the research cycle. Nowadays, a researcher is subject to stringent scrutiny by his peers at three points in that cycle:

1. when his research application or proposal is assessed (generally by senior figures within the field);

2. when he submits a paper for publication in a scientific journal, for presentation at a conference, or for inclusion in a book (submitted, under scrutiny);
3. after publication, by the entire research community; to use the terminology of A.D. de Groot, by the "scientific forum" [*wetenschappelijk forum*] (De Groot, 1961; De Groot, 1982).

Looking more closely at these three successive assessment points, we see an arena emerging consisting on one side of preliminary, subjective, speculative, and hypothesising concepts in the preparatory phase of the research proposal – a place where inspiration, imagination, daring, the urge to emulate, and other more subjective qualities are given scope within the *context of discovery* – and on the other of increasingly objectivised and reliable knowledge that is often scrutinised and criticised after publication (*context of justification*, Popper, 1972). This process of knowledge creation and knowledge output can be viewed as a funnel with a number of built-in filters from which the result of scientific enquiry emerges tried and tested (Haack, 2003: 197; Bauer, 1992: 45). The best results draw attention and remain. Ultimately, the result appears in scientific textbooks or disappears from them. The scientific forum does its work. Depending on the quality of the paper concerned, a contribution to science is discussed and cited (although a considerable number of published papers do not enjoy that privilege). These are the points of stringent assessment by science. But what is immediately apparent is that *between the first and second assessment points* – between the research proposal and publication – testing is left to the individual researcher or research group. This is therefore the phase when it is crucially important to exercise responsible data management. It is precisely in this phase that things can go wrong.

Ex-post corroboration of the accuracy and adequacy of the data and theoretical insights presented in the publication is possible if the experiment or study can be verified by other researchers using their own data and/or applying their own related or adapted hypotheses (verification or replication studies). Replication is the usual method of testing in the natural sciences and the life sciences.

This outline of the assessment points within the research cycle makes clear that quality control in a number of fields is based primarily on the later phases of the cycle, commencing with peer review of papers that have been submitted and, after their publication, the phase of free scientific discussion, controversy, and differences in interpretation.

The Committee has drawn the following conclusion from this survey of assessment points:

SECOND CONCLUSION

While the research cycle is subject to stringent and critical monitoring, there is one phase that is *relatively free of external review* and that leaves considerable scope for creativity. That phase can be found within the primary research process (i.e. the process of data acquisition, storage, organisation, processing, and accountability), in other words specifically after the start of the study and before the peer review. Depending on the field involved, this exploratory phase may represent a risk with respect to data management. The risks can arise due to shortcomings in m, although the presence and nature of those shortcomings vary from one field to another. A similar gap can be found in some fields after the end of the study, when the research data is archived. The absence of proper archiving makes it difficult to check the data at a later date.

Although the customary criticism and discussion within the scientific forum involves fewer risks, weaknesses can nevertheless be identified in the peer review system in a number of fields.

As noted in Section 2.3, there are differences within and between fields and not all fields fail to monitor researchers closely during the exploratory initial phase of research. The nature of the risks and shortcomings in monitoring can differ considerably, even within a single field. A discipline-specific survey is therefore needed to explore research practices within the various different disciplines. It is primarily a question of finding the right balance. If the "ex-post" assessment – by peers, peer reviews, and the scientific forum – are troublesome and not watertight, it is an obvious step to look more closely at the "ex-ante" phase, although the "ex-post" phase should also be tightened up. The more research reporting complies with the requirements of peer review and monitoring by the scientific forum, the less extra monitoring will be needed in the exploratory phase.

Analysis based on the responses to the Committee's seven questions can be useful in such a discipline-specific survey, although the survey will not make it possible to determine the frequency of good research and poor research or of effective and less effective monitoring practices.

2.5 Data management practices

Based on an analysis of the hearings – where follow-up questions were possible – and the written responses to the seven questions, we can comment as follows.

Are there concerns about how researchers deal with and manage data?

It was striking how often this question was answered with a resolute "no". The respondents generally felt great confidence in the scientific method, with the *nature of the field concerned* and the *nature of the data collected* both playing a major role. Respondents in the natural sciences (physics and astronomy, chemistry, the life sciences) and technical sciences referred to practices within their own disciplines, which combine stringent data monitoring with the possibility of verification and replication. Research that cannot be replicated or that is based on uncorroborated data will be shown up. Nature itself functions here as a kind of inborn peer reviewer. Even if the paper has successfully passed the test of peer review and has been published, it will receive little attention if its results cannot be replicated by other members of the research community (a fate suffered by some published articles). This leads to stringent self-censorship on the part of researchers. Nevertheless, a number of natural scientists note that this self-critical attitude is waning and that additional checks on the results of experiments (i.e. repeating the experiments a number of times) are becoming less frequent than in the past, partly because researchers are under such pressure to publish.

Other responses to this question referred to the nature of the research data. If that data is derived from a public source and is easily available to anyone, then monitoring can be so stringent that the likelihood of its being manipulated is considered small. Given the tough competition and the well-organised system of peer review, respondents are therefore confident about published data. They also had few concerns about data in fields in which researchers collect the data themselves but are required to submit a folder along with their paper – or at least make it available – accounting for the data.

In the case of large-scale research and research carried out by research institutes acting in partnership – often internationally – data acquisition and data management are regulated in protocols and collaboration agreements in almost all fields. There are hence few concerns regarding data (for example in astronomy, DNA banks, marine biology, international comparative social research).

The concerns that the respondents do express mainly involve individual researchers who collect, manage, and present the data on their own. Respondents say that in such cases, stricter checks by fellow researchers and/or better supervision are required.

Is data management monitored and are researchers properly supervised?

Where it is standard practice to store data centrally in the larger context of a research institute or research group, data management is generally adequately monitored. New or young researchers are inducted into this research culture and supervised by a supervisor or additional supervisor (senior researchers). The internal scientific attitude is then passed on by means of the ancient master-apprentice model or a group model, with the progress of the research and accountability being monitored at internal meetings convened to discuss the research and data management.

The respondents made various references in this connection to the large number of foreign researchers in the Netherlands whose primary training may not have involved training in data management practices. These researchers, who are generally PhD candidates, should be given a course at the start of their career to make up for that lack. One relatively frequent comment was that *all researchers just embarking on their career* do not have sufficient knowledge of advanced statistical methods to handle large and sometimes complex data files. This is partly because such methods are not covered in the Bachelor's degree component of the researcher's education. It is often wrongly assumed that new researchers have already acquired the typical tricks of the research trade during their previous study programme. The relevant gaps in their knowledge and experience can be filled in by having them take courses in data management and statistics.

However, in fields in which research and data management are often still the responsibility of single individuals (for example in patient-related PhD medical research and research in the social sciences and anthropology), respondents had observed shortcomings both in data management and in supervision. Data is then saved only on the researcher's own computer without his supervisors or colleagues knowing the details of how the raw data is processed. The respondents recognise that there is a greater risk in such individual research of inadequate data management; this indicates a need to view working within a team – something that is already typical of much scientific research – as the standard model.

Respondents generally had few complaints about research supervision; even the participating PhD candidates did not complain, although they sometimes felt that supervisors who had many other tasks did not *spend enough time* on supervision (for example research managers and medical specialists/research supervisors who also carry out clinical work). A few respondents also referred to their supervisor having only insufficient knowledge of the specific or specialist field of research concerned. Shortcomings of this kind can be obviated by the group research model referred to and praised by other respondents.

Peer review and validation of published data

Although the respondents expressed great confidence in the peer review procedure, which provides an evaluation of research results and comments at an early stage, several of them had serious criticisms concerning the way in which the procedure is implemented. They noted that journal editors sometimes asked authors to omit "superfluous" or "irrelevant" data, even though the author wished to account for that data. In such cases, researchers complied with the editors' instructions. In other cases, editors fail to comply with the rules and instructions that they themselves impose. The Levelt Committee refers to this in its interim report. The editorial boards of some highly respected journals failed to check for themselves how crucial elements of a study had been performed and how data collection was accounted for. This indicates that responsible data management is not solely the provenance of researchers and research groups; the monitoring bodies themselves must also be made aware of their duties. A number of disciplines, for example psychology, have long maintained the rule that the relevant data must always be submitted or made accessible, for example in a dataset or folder. That rule is not always complied with, however, nor are there sanctions for not doing so.

Where in the research cycle is there the greatest risk of something going wrong?

Respondents interpreted the term “wrong” in a number of different ways in their responses (i.e. the question was not formulated clearly enough, although it did produce valuable information). Some respondents interpreted “wrong” in the light of the Diederik Stapel case and assumed that it had to do with the possibility of fraud and deception. Many of those working in the natural sciences then stated that such deception did not in fact occur in their field or – as in the physics research conducted by Jan Hendrik Schön – that it was uncovered very quickly. Respondents pointed to the speedy refutation of the results of the cold fusion experiment, which nobody could verify, as evidence of the stringent verification system in the natural sciences. There is in fact little that can “go wrong”.

Other respondents confined themselves to the normal research cycle, pointing out that there were in fact risks in the initial phase of research. They stated that the greatest risk was not one of potential fraud but of a lack of expertise and incompetence, with supervisors not spending enough time overseeing young researchers, or not having enough knowledge of the research concerned. Data processing is difficult and often involves a great deal of calculation work, which in some disciplines is normally checked by colleagues and reviewers. Researchers in other disciplines, most notably within the social and behavioural sciences, are uninterested in habitually and systematically recalculating another researcher’s figures (see the interim report by the Levelt Committee). Respondents considered the points already identified as being susceptible to shortcomings high risk: a lack of statistical ability, “lone wolf” research or research that is not scrutinised sufficiently by others, and self-management of data, especially if the researcher leaves the organisation and nobody else knows how the data was collected or needs to be interpreted. They also noted that research by insufficiently qualified persons – for example Bachelor’s degree students carrying out research for PhD candidates – or small-scale research by very small research groups often lacks proper supervision and consequently generates results that cannot be accounted for later.

A few respondents referred to the danger of data frequently being “massaged” or processed for as long as it takes to confirm a given hypothesis (“trimming” and “cooking”). Respondents in one discipline noted that a certain data processing program automatically excluded “outliers” from the data file; as a result, the researchers, acting in good faith, did not investigate the outliers any further and presented examples of “cooked” data that were difficult to trace. In general, respondents were relatively satisfied with the way researchers handled their data; those who were not had more doubts about certain aspects of researchers’ professional expertise than about their intentions to report on honest and reliable research.

What happens to the research data after the research concludes? Is data archived and made available?

When it comes to long-term data storage, data availability and data archiving, practices differ enormously from one discipline to the next. In some fields, for example physics and chemistry, such matters have are meticulously regulated and good practices have grown up whereby researchers can retrieve research data for verification purposes or can keep it available for themselves and others so it can be utilised again at a later date. In astronomy, the overwhelming majority of data is properly archived and, after a certain period of time, is available to every researcher.

In other fields, data archiving and data sharing are still in their infancy. Where data is collected on a large scale – for example in the Dutch Parliamentary Electoral Studies (NKO) in the field of political science – there are satisfactory arrangements for making the data subsequently available to other researchers. In the case of longitudinal studies – in sociology, developmental psychology, and educational theory – researchers have gained positive experience managing and archiving large databases. The Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO) set up Data Archiving and Network Services (DANS) in order to promote long-term access to digital research data. With that object in mind, DANS maintains digital data archives in various scientific fields.

A few respondents noted that research data could *not* be made available for verification by other researchers because the relevant data files had to be paid for.

Others also noted that in research data in the behavioural and social sciences and humanities was not (yet) made easily available to fellow researchers. That meant that reanalysis or replication could not take place – assuming that this is was in fact standard practice. The data concerned should be available for at least five years. Some funding organisations have not yet established any specific requirements regarding access to collections of data that they have financed or regarding the quality of digital data repositories. A few disciplines are in fact developing such professional repositories, however, in collaboration with DANS. Part of their motivation is to comply with the integrated European Framework for Audit and Certification of Digital Repositories.

The survey questions encouraged an assorted group of willing researchers to provide reasoned responses that offer a valuable impression of the strengths and weaknesses of research data management in the various fields. Readers should be warned, however: the survey is nothing more than an impression. Its findings were not obtained by systematically examining a random selection of all researchers, studies, research groups, and scientific fields. Given the current "loaded" nature of the topic, it is possible that the respondents/correspondents gave socially acceptable answers and painted too rosy a picture of practices in their field. Only an in-depth, representative investigation of research practices in the various fields can obviate these limitations of the Committee's analysis (see Sections 2.8.2 and 3.6). Just how such an investigation should be initiated and carried out is primarily a matter for the scientific organisations TNO, NWO, VSNU, and KNAW to discuss.

Given the unmistakable variety revealed by the survey, the Committee draws the following conclusion:

THIRD CONCLUSION

With so many differences between and within scientific fields, the best approach would be to consider the risks, shortcomings in monitoring, and options for improving monitoring in the exploratory phase of the research cycle *for each separate discipline*. Gaps in ,monitoring can also best *be identified* for each separate discipline. The various disciplines can learn from the good practices of other disciplines and – if necessary and possible – take over various monitoring mechanisms from those disciplines (for example keeping diaries and logbooks or lab journals, a folder accounting for data, carrying out research within teams, and ensuring peer pressure prior to the peer review phase).

The Committee finds it not only advisable but also unavoidable and consequently necessary to refer back to the various scientific fields and separate disciplines when developing instructions for responsible data management. There and only there can possible and necessary improvements be identified, along with the most suitable methods. The professional associations and societies associated with the scientific fields have as much of a role to play in this regard as the research institutes and the universities. After all, responsibility for the quality of research does not lie solely with the individual researcher or research group, or solely with the organisation that bears formal responsibility, i.e. the university, but just as much with the research community (or communities) itself. But what is most striking about that community – when compared to the universities and research institutes, which are responsible for appointments, salaries, etc. – is its informality. The research community consists in part of countless informal circles and networks. Local, national, and international networks ensure constant communication between researchers. Within these networks, researchers provide informal assessments of the quality of each other's work and research; they arrange new research projects, and extend prestigious invitations to speak at conferences or to join the boards of scientific associations (including international ones). They are also where important positions are allocated and invitations are extended to join the editorial and advisory boards of scientific journals. In other words, formal organisations such as universities and research institutions – no matter how much they join forces – can never get through to the heart of research if the informal networks of researchers do not cooperate. As already pointed out, the research community consists of countless informal circles and networks. There is constant communication between researchers via local, national, and international networks. Every researcher is also a peer, i.e. someone who assesses the work of other researchers in his field, as a reviewer, an editor, a conference or seminar organiser*, and as a member of doctorate or external review committees. It is within these scientific communities that the culture of a discipline is created, together with its values and standards for* *acceptable and unacceptable behaviour. The universities have hardly any influence on this. This situation imposes a heavy burden of responsibility on these scientific communities and their members; they must be constantly vigilant concerning the standards and values for data management within their particular field.*

The Committee's recommendation is simple and to the point, then: do not attempt to develop an overall policy but remind every scientific field of its duty to conduct research responsibly and monitor whether and how the various fields fulfil that duty.

2.6 Corresponding features of good practices

Based on a number of criteria, for example the scale of the research concerned, it is possible – regardless of the discipline – to draw a number of conclusions that transcend the individual disciplines and can consequently be used to reinforce research practices where that is considered necessary or desirable.

One of the first insights that emerged from the survey is that the extent to which research is verified depends largely on the scale of that research. The greatest risk of something going wrong is in individual research, where data is collected, stored, and managed by only a single researcher who may receive only superficial supervision. The risk here concerns both responsible conduct and the potential for questionable research practices. This does not mean, however, that we should prohibit or discourage individual or isolated research; as a matter of fact, such research can make a creative contribution to science. The point is to reduce the potential risks by providing proper supervision and verification, in the form of discussions with others and accountability after the study has been concluded.

At the other end of the spectrum is large-scale research, which – particularly when conducted in an international context – generally (and indeed almost exclusively) involves stringent monitoring of data collection, management, and subsequent archiving. The numerous partners involved in the research "keep an eye on" one another, especially because they all depend on one another's data collections and the quality of those collections. This does not mean, however, that all research should be performed and organised in this way. The Committee considers it extremely important for small-scale, creative, and unusual research to be carried out and to remain possible; what is needed is to improve verification and supervision for such cases.

Good practices also emerge from the picture painted by the various fields. Research data is properly managed when the nature of the discipline itself offers verification methods that the researchers and the research culture can emulate. Other disciplines have developed their own good practices, for example having data available or making it so, keeping lab journals (and more generally: maintaining research journals explaining how a particular study has proceeded) and promote reciprocal data checking so that comprehensive internal controls are in place in the primary research process. In some fields, the data itself is such that other researchers in the field will quickly detect any manipulation; here, self-management plays a decisive role in the way researchers deal with data, and the scientific forum functions effectively in its self-monitoring and self-correcting role. It is only where such elements are missing, either in the research practices or the research *culture* of the discipline concerned, that there is reason to learn from the good practices of other disciplines and to take on board elements of those good practices where possible.

FOURTH CONCLUSION

Good data management practices that are already established within a number of scientific fields can, where necessary and applicable, be introduced within other fields. Examples include arrangements and protocols in international research; established procedures for collecting, managing, storing, and using research data in large-scale research; and the inclusion of external monitoring mechanisms in research that is frequently carried out by individual researchers working in isolation, owing to the nature of the subject.

Potential shortcomings of this kind of research can be tackled by improving research practices (see the next section). However, the *overall* picture that emerges from the survey does not really give cause for concern. Emerging research practices indicate that research is being carried out properly and that good practices are being taken on board: more reciprocal monitoring, more systematic verification in the shape of protocols, more data sets submitted with articles, etc. What is lacking – and has in fact been lacking for a long time – is the awareness that there are in fact benefits to having a responsibly designed data collection system and to the bodies responsible pursuing a policy aimed at inculcating and watching over that awareness in the workplace. In fact, there was complete faith in the time-honoured image of science as an autonomous system that monitored its own practices. The outside world was shocked, however, by a number of spectacular examples of scientific misconduct and began to wonder whether that misconduct was perhaps representative of science in general. That is definitely not the case, but proof that it is not can only be found in the good practices in each scientific field.

2.7 The Committee's main recommendation

Given the four conclusions above, the Committee's main recommendation is that the world of science should itself make its existing research practices – including those that require improvement – transparent to the outside world and to society in general. Improvements should not take the shape of regulation or supervision imposed from above or from outside science, but should be implemented and accounted for by the various scientific disciplines themselves. In particular, of the many improvements that may be possible, each discipline should undertake those most appropriate to it. The motto for all the best practices suggested above should be "if the shoe fits, wear it". It is not possible to develop a strategy that can be implemented in all scientific disciplines simultaneously – the differences between them are simply too great. While this recommendation does permit freedom of choice, it is by no means a recommendation that is free of obligation. Improvements can be made in every discipline, and each discipline should explore and implement such improvements in its own research.

2.8 Recommendations

2.8.1 Divergent and shared responsibility

Three specific aspects have consistently emerged from our analysis of the issue of responsible data management in research:

1. the *individual* responsibility of the researcher;
2. the *institutional* context within which the research is carried out;
3. the *informal* networks that play a vital role within the research community (or communities).

The important thing is to achieve a stable balance between these aspects. The responsibility borne by the formal organisations that facilitate and fund research does not cancel out each researcher's individual responsibility for the quality of his own research. Research is basically an exploration, and the creative phases in the research process are impervious to central management or regulation, and so they should be. Research groups themselves and a culture of responsible conduct at research institutes have a more direct influence on actual research, and they are closer to it. In this respect, they bear more responsibility for ensuring good quality research. The informal scientific communities play a key role by conducting peer reviews and by verifying scientific accuracy; however, they are to a large extent separate from the universities and funding bodies, which in their turn neither control nor are responsible for journal editors and academic publishers.

Although this broad spectrum of scientific practices is undergoing rapid change – with open access to publications and data being familiar examples – the Committee considers that, where necessary or desirable, all these different groups and organisations can be asked to coordinate simultaneous improvements in research data management. The Committee's recommendations on responsible research conduct therefore address multiple levels.

Recommendations for improving research data management should be taken to heart at various organisational levels, levels that are fairly easy to identify.

1. First of all, there is the researcher himself. He is responsible – at all times and throughout the research process – for responsible data management. The fact that he may be working as part of a team does not excuse him. However, he should be encouraged to internalise this responsibility by the environment in which he works.

2. The second level is that of the workplace where the research itself is carried out, and where researchers assist and supervise one another. This therefore encompasses individual researchers, the research group, and research and researcher supervision by PhD supervisors and other senior figures. An open scientific culture – meaning one in which researchers discuss their work and allow it to be observed and monitored by others – is of the utmost importance.

3. The third level is that of the separate scientific disciplines. They and their professional associations have created the scientific culture of their own journals and conferences. They maintain that culture even if it differs from one discipline to another. It is therefore also up to these disciplines to preserve this culture as undisputedly scientific in nature, and to put things right when it becomes necessary. This may mean improving the system of peer review when flaws in verification become apparent, or raising awareness within professional organisations of the scientific standards and values and of conduct that may undermine those standards and values.

   The disciplines can bring about changes in the pressure to publish, for example, or in the practice of co-authorships, where and if such practices have a negative impact on the quality of the research. This would include fellow researchers jointly ensuring the transparent, open access archiving of data, with the necessary monitoring of research data being regarded as a fixed and normal part of their professional work.

4. The fourth level consists of the official bodies that control and are responsible for scientific research and individual researchers. This includes: research institutes and their directors (institutes where a large number of researchers work and where the topic of "data management" can be made an intrinsic component of performance appraisal interviews); faculties and faculty directors of research; and the responsible bodies within the universities and university medical centres (management boards and executive boards at UMCs). Each of these bodies – in its own way and within its own remit – can promote a good research climate in which due attention is given to responsible data management. They can no longer ignore the topic of data management when conducting reviews of research.

5. Finally, the overall task of umbrella organisations such as the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Association of Universities in the Netherlands (VSNU) is to support the good intentions of the researchers and the disciplines and to encourage the disciplines to take action.

The Committee's recommendations for each individual level are given below. The Committee considers it important to work with these recommendations for a time before assessing their effects; because practices differ so much within and between disciplines, not every recommendation will apply to each and every one of them. Only after the disciplines have worked with the recommendations for a number of years will they have enough experience to determine whether doing so has led to real improvements. The umbrella organisations should keep close track of how the various recommendations are implemented within the various disciplines, , however. The recent wake-up call that alerted the scientific world as a whole to the need to defend the reliability of science must continue to ring out loudly and should not be quickly muted.

The Committee's recommendations primarily describe the most *desirable* situation in scientific research. If that situation is already established practice in certain disciplines, it will be enough to maintain it and perhaps raise awareness. Improvements are necessary where the most desirable situation has not – or not yet – been achieved. It follows from the Committee's main recommendation that each discipline must consider for itself whether and where improvements are necessary.

2.8.2 Specific recommendations

Re 1: The researcher himself

Every researcher must be aware of the stringent requirements imposed on the performance of scientific research: researchers must act responsibly and meticulously, be ready to account for the main steps taken in the course of the research, keep an open mind towards negative results that refute their hypothesis and expectations, and be willing to accept criticism. A scientific mind-set also means they must strive for objectivity, ideological neutrality, and independence. There are few studies examining how researchers – both experienced researchers and those just embarking on their career – should internalise these values. It has always been assumed that these values were transmitted by means of role models and the master-apprentice model. The changing university culture – many of the respondents referred to the rising pressure to achieve and to publish – and the growing importance to society of research results make it questionable whether this traditional means of transmission is sufficient.

The Committee considers that all researchers must deliberately cultivate and demonstrate a scientific mind-set, including the realisation that they themselves bear primary responsibility for responsible research conduct and responsible data collection, storage, management, processing and reporting. They may well share that responsibility with other researchers and with their supervisors, but that does not excuse them from shouldering their own share of it. It is up to the research institutions to create and maintain a climate in which such responsible research conduct can flourish.

Re 2: The research group and the research institute

Greater peer pressure

One of the Committee's most important suggestions concerns rigorous mutual vigilance on the part of researchers and research groups. Even before papers are peer reviewed, those most closely concerned – colleagues, research coordinators, PhD supervisors – should apply the necessary peer pressure throughout the whole course of the research. It is not enough to simply "muddle through" with data and statistical analysis of that data in the initial phases of the research. In one of his science columns in the Dutch newspaper *NRC Handelsblad*, Piet Borst referred very appropriately to the "web of social control" (Borst, 2002). Pressure from one's colleagues and superiors can have both good and bad consequences. The emphasis tends to be on the negative consequences – i.e. the "one rotten apple" theory of social influence – but in a properly organised environment, informal social control can also have a positive effect. This means that researchers must pay attention to the day-to-day work of their counterparts both within and beyond their own research group. Positive peer pressure is also promoted if the PhD supervisor or research coordinator visits the workplace regularly to ask about progress and to talk to the researcher/researchers, or visits the sites where research data is being collected. He or she is then more likely to spot "something wrong" with the data or with the procedures being used. By taking an interest, he or she can promote an atmosphere of openness, not merely in terms of intellectual debate but also concerning the practical aspects of data collection.

Greater social control by peers is a modern version of the traditional master-apprentice relationship, with the scientific mind-set and research ethics being transmitted – informally – from one generation to next, generally from the supervisor to the PhD candidate or assistant. This master-apprentice model can still be valuable if it is adapted to contemporary circumstances. Most research nowadays is not carried out by individuals; it is generally a team effort involving not only the primary researchers but also the experts and supervisors. A PhD candidate often has two or three supervisors, with all the associated advantages ("two heads are better than one", expert advice and monitoring). The one-to-one relationship also has its disadvantages (unwanted pressure), as does double or triple supervision (contradictory advice, supervision neglected because one supervisor believes the other is undertaking it), and these must be avoided as far as possible. Horizontal peer pressure and the vertical master-apprentice relationship can complement each other effectively.

Systematic study of data management practices

The Committee recognises the need for a *systematic study* of research practices in the various scientific fields. It suggests investigating, a number of *randomly* selected research projects, either completed or ongoing, at a large number of research institutes over the next four years, along with the researchers involved. Questions could include: how was the research set-up designed and by whom; how is the necessary data (both raw and processed) selected and by whom; how is that data stored and shared; what were the most important and the most difficult research decisions and how were they taken; and how are the results of the research being substantiated and reported? We could learn a great deal from such a study. Complete openness about the "life cycle" of an actual piece of research would not only clarify how research data is dealt with but also provide important insights into how that is done, and help identify crucial differences between scientific disciplines and the way in which they solve or cope with certain recurring data management problems. The study would also focus on technical aspects of data management specific to a given discipline, leading to improvements.

The Committee believes there is a need for a verified scientific insight such as this. It will not only remove any remaining doubts about the reliability of research data but also comply with the precautionary principle that should prevail at formal organisations (the universities, the Academy, NWO), namely to "be prepared". The insights gained over this relatively short period could be used to improve vigilance with respect to responsible research practices.

Another way to efficiently organise critical supervision was proposed by the researcher Prof. Martijn Katan at the Academy's research data meeting (December 2011). Prof. Katan proposed random inspections at research sites, followed up by an interview in which the researcher/researchers would be asked how the data had been acquired, utilised, and analysed. Prof. Katan proposed a kind of "flying squad" that would visit researchers and research institutes unannounced in order to see how data was recorded and analysed and whether it was possible to replicate the published results using that data. The Committee considers Prof. Katan's idea for such a system of random inspections of researchers and research institutes to be worthwhile.

The Committee's proposal to conduct a *study* – also randomised – on research data is inspired by Prof. Katan's suggestion and is in fact a variation on his idea. If the emphasis is on the need for a sudden and unexpected *inspection* of research practices, then it will be sufficient to set up a "flying squad". In the view of the Committee, however, it will take an actual study to gain a greater understanding of the various research and data management cultures and of the technical options and problems associated with the latter in various scientific disciplines.

Something that the Committee considers just as important is for scientific organisations at all levels to hold themselves accountable for honest and responsible data management: honest in the sense of accurate calculations, precise work, and – above all else – no doctoring of data (see Section 3 of the present Advisory Report); responsible in the sense that the steps and decisions taken *at crucial moments* during the research and its results can be accounted for (note that this does not mean accounting at *all times* for every detail of a study or for every *scribbled note*; that would make research virtually impossible).

Everyone involved in the research process, regardless of their status as a junior or senior researcher, must accept their share of the responsibility for this. Individual researchers and research groups must monitor and question one another critically, for example by engaging in regular "peer coaching". Supervisors (including PhD supervisors) and research coordinators must take responsibility for actual supervision and not restrict themselves to "remote" supervision.

Re 3: The scientific disciplines

Journal editors must ensure accountability for research results for each manuscript submitted, for example by requiring documentation to be available on the data and related syntax for statistical routines. Space should be made available in publications for "negative" research results and for verification or replication studies. Editors should be given access to the underlying data if they so wish. As a general rule, when an article appears in a reviewed journal, the underlying research data should be available to other researchers. Rules regarding co-authorship should be reconsidered and where necessary tightened up so that co-authors are required to be familiar enough with the research data to allow them to share responsibility for it. The Committee is aware that the rules regarding co-authorship differ from one field to the next, but this does not mean that existing practices do not need to be reconsidered.

Professional associations – many of which organise scientific conferences – must subject their discipline's research priorities and remuneration structure to critical re-evaluation. Although universities and the Academy do not maintain formal relationships with these professional associations and societies, it is nevertheless up to them to require these organisations to exercise vigilance with respect to reliable and responsible scientific research. After all, the wake-up call applies just as much to the informal as to the formal scientific domain.

Re 4 and 5: Responsible bodies

Research institutes should support the research process throughout the research period, for example by ensuring that a data manager and statistics consultant are available, who would not only reconfirm the data at the end of the research but also advise on the research design and implementation and on data storage and processing. Deans and directors of research institutes can include the topic of responsible data management in their annual performance appraisal interviews with all researchers, whether they are just starting their career or are more experienced. Regular (e.g. annual) study days to discuss data management and data processing problems can encourage researchers to remain conscious of the need for responsible research.

Management and executive boards at universities and university medical centres (UMCs) should see to it that the primary research process is given the proper amount of attention, with time and financial resources being made available within the organisation. Similar to the performance appraisal interviews at research institutions, management and executive boards should include responsible data management as a fixed component of the Standard Evaluation Protocol (SEP) for research or institutional evaluations. This responsibility rests primarily with the bodies that are drawing up the SEP (VSNU, NWO, and Academy).

Management and executive boards must understand and accept that the process of systematically improving research monitoring will cost money and take up researchers' time.

3. INTEGRITY IN SCIENTIFIC RESEARCH

3.1 Integrity as a guiding value in science

Integrity is becoming an increasingly "hot topic" in a number of sectors of civil society, and the discussion concerning scientific integrity is part of that trend. The 1986 "Baltimore affair" – in which the spotlight was on alleged manipulation of research data – marked a turning point (Kevles 1998; Heilbron 2005, 30–32; Marks 2009: 185–189; Goodstein 2010: 61–64). The case, which was only concluded in 1996, led to anxious discussion within an American Congressional subcommittee and to the founding in 1994 of the official *US Office of Research Integrity.* This was preceded two years earlier by a report by the US *National Academy of Sciences* and by a number of separate studies. Scientific integrity is an issue of concern to all fields of science.

In the Netherlands, the Academy, NWO, VSNU, and the universities have jointly recognised the importance of responsible research conduct and scientific integrity. In October 1995, the Academy, NWO, and the VSNU published a memorandum on scientific misconduct [*Notitie inzake wetenschappelijk wangedrag*]. This was followed in 2001 by a memorandum on scientific integrity [*Notitie wetenschappelijke integriteit*]; that same year, the Academy published the booklet *Scientific Research: Dilemmas and Temptations* [*Wetenschappelijk onderzoek: dilemma's en verleidingen*] (Heilbron 2000; 2005). In May 2003, these same organisations set up an independent *National Board for Scientific Integrity* (LOWI) along the lines of the *US Office of Research Integrity,* following the example of comparable bodies in other countries. Since 1 January 2005, all researchers at Dutch universities have been required to comply with the *Netherlands Code of Conduct for Scientific Practice*, which also applies to the various Academy institutes. The NWO institutes also have a code (NWO, 2005). Since 2004, numerous universities and university medical centres have introduced their own regulations and code of conduct, with the final ones doing so in 2012.

The growing interest in scientific integrity can also been seen in a number of important international reports published in recent years, in particular *Ensuring the Integrity, Accessibility and Stewardship of Research Data in the Digital Age* by the United States National Academy of Sciences (2009) and the OECD's report *Best Practices for Ensuring Scientific Integrity* (2008). International codes of conduct have also been drawn up, for example the *Singapore Statement on Research Integrity* (2012) and the *European Code of Conduct for Research Integrity* (2011); the latter is the result of collaboration between the European Science Foundation (ESF) and the federation of All European Academies (ALLEA). The Code has been adopted by ALLEA and the European Science Foundation (ESF) and made binding on their members.

Scientific integrity is part of the ethics of scientific practice. These ethics cover more than simply integrity and concern more issues (including the treatment of patients, human test subjects and laboratory animals, and relationships with clients). The meaning of integrity is more specific; it can be viewed as the societal value associated with specific positions within society. Specific standards of conduct have emerged for these positions and professional groups, which vary from one sector to the next. Whether an action is identified as misconduct or as a contravention of standards is thus context-bound. Public servants, for example, must not accept money or favours in return for the service that they provide or for decisions that they must take, nor must they use their position to extend favours to others (Huberts, 2005; Lamboo, 2005). Doctors have a universal duty to help anybody whose life is in danger and they also have a duty to observe professional secrecy (Kirkels, 2004). Lawyers are also subject to a professional obligation of secrecy, and must remain independent vis-à-vis their clients. Judges must remain impartial at all times and journalists must be truthful in their reporting. In the commercial world, conflicts of interests (financial or otherwise) and misuse of insider information must be avoided. The rules concerning the use of guile and deception in transactions are less strict, however, because purchasers also have a duty to proceed with caution when buying a house, a second-hand car, or a share portfolio. Bribery and backhanders are not permitted in our culture, however (Rippen, 2006: 61–90). These examples show that integrity always involves *specific standards* and characteristic *values* that are regarded as guidelines for those holding certain positions. These characteristic, guiding values differ from one social sector to another, as can only be expected in an open society differentiated by tasks and activities.

The *guiding values* of science differ from those of other sectors of society. In this approach, scientific integrity is interpreted in the narrower sense. It does not cover failure to comply with *general* legal requirements (for example the prohibition on theft or blackmail). We are not obliged by society *not* to invent something or *not* to write down what we have invented, whereas in the world of science inventing data is a serious contravention of the standard of responsible scientific behaviour. It is normal to invent stories in novels and other fiction, whereas in science doing so contravenes the relevant standard. Copying is permitted within numerous social institutions, for example within the family, when presenting news in press releases, and in church sermons. A government minister giving a speech may use a text written by one of his civil servants. In the law, copying acts of parliament and case law rulings verbatim is both justified and desirable, but it is not permitted in the world at science (at least not unless the source is indicated). In the Middle Ages, monks preserved and passed on our culture by copying out texts; nowadays, however, copying is prohibited at research institutes and laboratories – today's scientific "monasteries" – without proper source citations. Science requires stricter and more precise standards of honesty than does everyday life. Scientific integrity is therefore a specific standard that is linked to the researcher's position – a standard *set for his or her own behaviour*: this is how we conduct ourselves in this profession.

It was the British literary scholar and novelist C.S. Lewis who gave the shortest definition of "integrity" as a value: "*doing the right thing even when no one is watching".* In Lewis's view, integrity is the conscientious judgement of *one's own behaviour* (Lewis, 1962: 146–147; see also: Lewis, 1960: 191–196). Lewis describes precisely what lies at the core of integrity, namely *personal responsibility.* In another definition of integrity, the British philosopher Bernard Williams echoes Lewis in emphasising the responsibility that someone feels for their own conduct, leaving aside utility and the conduct of others or the requirements that others impose (Williams, 1973: 116). What is involved is a *commitment* that we enter into with the values that we have chosen. Those values give us our own moral identity and our conduct is in line with them. It is true that Williams' definition of integrity is a general one, but it can be used effectively to produce a definition of scientific integrity: a personally chosen commitment to the special values of science.

Other authors also emphasise personal responsibility and conviction, answering for what one believes and does (Carter, 1996). In its definition of integrity, the *Oxford English Dictionary* emphasises the absence of its contrary: "*freedom from moral corruption*"; it defines integrity in terms of honesty, sincerity, and truthfulness. "Scientists have intellectual integrity in so far as they strive to follow the highest standards of evidence and reasoning in their quest to obtain knowledge and avoid ignorance" (Resnik, 1998: 84). One chooses science and the specific values of science. The result of integrity is, above all else, mutual trust. If scientists are unable to trust one another, scientific progress is impossible; it goes without saying that they must be able to build on one another's work.

Emphasising integrity as an individual trait, however, distracts us from the structural component that is also part and parcel of integrity. After all, organisations responsible for integrity – the police and judicial authorities, the Bar Association, banks, churches, universities – need to consider how they can prevent a situation where "no one is watching" from arising too often. Just as with responsible research conduct (see Section 2), integrity involves striking the right balance between individual responsibility and the responsibility that organisations bear for the circumstances in which individuals can and do work.

The idea of scientific integrity as a personal, internalised standard that is also supported by one's work environment leads us to question how that standard is internalised and how we can encourage it. Personal standards develop through a process of reciprocal influence and exemplary behaviour, with the human ability to internalise cultural influences playing a crucial role. People express in their conduct the influences that they have felt and have internalised. They thus create role models for themselves. In a scientific context, such role models developed chiefly within the master-apprentice system and, today, in the relationship between PhD supervisors and their candidates. This internalisation process – which covers much more than social control and the threat of sanctions – is crucial to the development of a scientific attitude and scientific integrity. In this context, the research community has a clear duty to promote the internalisation of the standards and values of science (see also Section 3.6 and Section 4).

3.2 Scientific fraud as a warning

Although scientific fraud has played a subordinate role in the history of modern science, it does attract attention. Researchers in a number of disciplines have attempted to deceive their colleagues, the scientific world, and the public by fabricating data. Notorious cases include that of the "Piltdown Man", a set of faked anthropological remains supposedly excavated in England in 1912 and said to be those of an early human; the case of the biologist Paul Kammerer (Vienna 1910–1923), in which unusual markings on midwife toads turned out to be Indian ink injections; the Cyril Burt case, involving falsified research data on the intelligence of twins reared apart (London 1954–1966); and experimental skin transplants carried out on mice by the dermatologist William Summerlin at the Sloan-Kettering Cancer Centre (New York 1974); the fact that some of the animals had been coloured using a felt-tip pen came to light when Summerlin's lab assistant cleaned them with alcohol.

In two recent cases of scientific fraud – involving physicist Jan Hendrik Schön (2000) and Korean stem cell researcher Hwang Woo-Suk (2008) – the perpetrators were regarded as "brilliant" researchers. When fraud is uncovered, the public's confidence in science's self-refining ability is reinforced, but people are still left wondering how such cases could have occurred in the first place (Broad and Wade, 1983, Heilbron, 2005; Marks, 2009; Goodstein, 2010).

In 1993, the Dutch journalist Frank Van Kolfschooten followed the example of science journalists Broad and Wade by producing an overview of fraud, deception, and plagiarism in the Netherlands (Van Kolfschooten, 1993, 1996, second edition). Van Kolfschooten related how bacteriological research on tumours in fish supposedly carried out between 1952 and 1962 by Anthonie Stolk, a professor at Amsterdam's VU University, had never in fact been performed. Stolk also reported on medical-anthropological research that he claimed to have conducted among indigenous peoples in Africa, without ever having actually been there (Van Kolfschooten 1996: 122–136).

Van Kolfschooten's book appeared just when the Diekstra and Buck cases were in the limelight in the Netherlands, and it inspired the scientific world (the Academy) to take the first steps towards combatting scientific fraud. In light of the Diekstra case, the Faculty of Social Sciences at Leiden University developed a protocol identifying the ethical principles for scientific research, together with guidelines for action to be taken in cases of alleged scientific misconduct (Protocol 1998).

The fraud committed by the social psychologist Diederik Stapel at Tilburg University is another case of exceptional deception by a highly imaginative "lone wolf". It is unique in the same way as the cases outlined above, and in no way representative of "normal science". The discovery of Stapel's fraud did act as a wake-up call, however, and it also revived the issue of scientific integrity, both within political circles and in the world of science.

Has scientific fraud always existed? In his book *Reflections on the Decline of Science in England, and on Some of Its Causes*, the mathematician Charles Babbage, pioneer of computer science, wrote about fraud as a contemporary phenomenon as far back as 1830, specifically mentioning two types: "trimming" and "cooking":

> *"*
>
> *Trimming consists of clipping of little bits here and there from those observations which differ most in excess from the mean, and in sticking them on to those which are too small." (Babbage, 1830, ed. by Hyman 1989: 122 ff.; quoted in Broad and Wade, 1983: 29–30).*

Babbage disapproved of "trimming" but understood that it did occur and that it was in any case less reprehensible than other types of fraud:

> *"The reason of this is that the average given by the observations of the trimmer is the same, whether they are trimmed or untrimmed. His object is to gain a reputation for extreme accuracy in making observations; but from respect for truth, or from prudent foresights, he does not distort the position of the fact he gets from nature." (op. cit.).*

Babbage considered "cooking" to be much worse:

> *"...to give ordinary observations the appearance and character of those of the highest degree of accuracy. One of its numerous processes is to make multitudes of observations, and out of these to select those only which agree, or very nearly agree. If a hundred observations are made, the cook must be very unlucky if he cannot pick out fifteen or twenty which will do for serving up." (op. cit.).*

But what Babbage considered most reprehensible was:

> *"...the scientist who pulls numbers out of thin air. The forger is one who, wishing to acquire a reputation for science, records observations which he never made. Fortunately instances of the occurrence of forging are rare." (Broad and Wade, 1983: 30).*

The quotation about the "cook", whose claims are not supported by the meal he prepares, is a good analogy for a certain way of manipulating research data, namely revealing only positive results – something that is a constant temptation for researchers. This actually means that the researcher provides information selectively, with the problem being not so much his selection of raw data as the fact that he systematically omits unwelcome data. It should be noted, however, that focusing selectively on certain aspects of the data can in fact produce new insights, for example by meticulously investigating one or more "outliers". In such cases, however, the researcher must make clear that he has selected specific data, how he has done so, and to what extent.

3.3 A spectrum of scientific misconduct

In what circumstances has scientific integrity been contravened, and what criteria apply to such contravention? The appropriate term here is "scientific misconduct", and classic examples are not difficult to find:

- fraud with research data: inventing, making up, or fabricating data oneself;
- manipulating or misrepresenting data, for example by omitting negative outcomes (what Babbage refers to as "cooking and trimming");
- plagiarism, both literally but also by "stealing" or using ideas without crediting their source.

This is the familiar trio: fabrication, falsification and plagiarism (referred to as "FFP" in international discussion). Resnik adds *fraud*, a term referring to its opposite standard of *honesty* in science (Resnik, 1998: 74–80). Fraud is a broader concept than fabrication or falsification because it also includes omission, for example when a researcher says that he has tested something without actually having done so.

There will be little disagreement about these three types of misconduct; after all, they strike at the very heart of the ethos of scientific research and the pursuit of science, which is based on openness, honesty, reliability, and verifiability. But when can an action be regarded as plagiarism? Is it plagiarism if a researcher copies just one or a handful of lines from someone else's work – without that person's consent – or does he need to have copied extensive passages, or half an article, or a whole article? Can a researcher also appropriate someone else's scientific *ideas*? When is there clear evidence of data manipulation? Is it only when certain outcomes (unwelcome or chance results) are omitted or can we say that data is already being manipulated when a researcher, driven by his own or some idiosyncratic theory, is biased in the way he interprets the outcomes – something that he might do unconsciously? Has scientific integrity already been compromised if a researcher makes a mistake, for example by applying the wrong methodology in his study?

The *Netherlands Code of Conduct for Scientific Practice* (2004) makes clear that scientific integrity is contravened in the case of any one of the above types of misconduct.

These basic rules were codified in 2004, although they have always existed and formed the implicit basis for modern scientific practice. In the view of the Committee, they do not need to be tightened up. They are already clear and straightforward enough: “thou shalt not plagiarise” and “thou shalt not produce or publish false data” (i.e. fabricate or manipulate data). What is more important is the degree to which all researchers are aware of the rules and comply with them in actual research practice. There are naturally borderline cases in everyday scientific practice, and a large grey area that is not described as such in the 2004 version of the *Code of Conduct*.

The 2012 revision of the *Code of Conduct* summarises the main contraventions of scientific integrity, referring to:

a. falsification of data;
b. inputting fictitious data;
c. secretly rejecting research results;
d. deliberately misusing methods (including statistical methods);
e. deliberately misinterpreting results;
f. plagiarising other people’s publications and results (or parts of them);
g. wrongly presenting oneself as a co-author;
h. deliberately ignoring or failing to credit contributions by other authors;
i. culpable lack of scrupulousness when carrying out research.

(VSNU, *Netherlands Code of Conduct for Scientific Practice* 2012).

In cases of serious contravention of scientific integrity, it is appropriate for the employer to impose sanctions. In cases of scientific misconduct, however, such sanctions are not enough. Formally speaking, the *Netherlands Code of Conduct for Scientific Practice* (2004; 2012) regulates the conduct of university researchers. It is an important document, and some universities – for example Erasmus University Rotterdam and Leiden University – issue a copy to every new PhD candidate and staff member.

But is all *undesirable* conduct also conduct that *lacks integrity*? What is the difference between misconduct and undesirable conduct? What kind of standards and systems apply when we refer to a contravention of the basic rules of scientific practice or a failure to comply with the standards set out in the *Netherlands Code of Conduct for Scientific Practice*? Paradoxically enough, serious contraventions make the system of standards clearer. The Diederik Stapel case offers a good example of an idea first put forward by the sociologist Émile Durkheim at the end of the nineteenth century: publicly transgressing a central group standard promotes group cohesion precisely because the standard is once more brought to the group’s attention and contraventions meet with strong disapproval (Durkheim, 1893: 69). The contravention serves to reinforce the original group standard and help the group internalise it. This is a somewhat ironic but predictable consequence of noticeably deviant behaviour that meets with the community’s disapproval. Like previous fraud scandals in other countries, the

Stapel case has had a comparable effect on science in the Netherlands. People were uneasy and shaken up, and immediately felt a pressing need to tighten up the rules again.

In philosophy, a distinction is made (based on the work of Wittgenstein) between *game rules* and *goal rules,* in other words between the basic rules of the game on one hand and rules of thumb for achieving the goal of the game on the other, similar to the difference in chess between invalid and unwise moves. To break the basic rules of scientific practice is to contravene an absolute prohibition: “thou shalt not commit fraud” and “thou shalt not plagiarise”. The authority of science must be preserved by making a clear distinction in this case. In other words, these are objectivised minimum standards of scientific practice; here, science functions as a social institution that generates new knowledge by applying controlled methods. It is a serious matter to contravene these rules; the integrity of science is at stake. Such conduct merits a “red card”, so to speak; it is a breach of the *game rules*. In other words, it is no longer science; science itself is harmed.

Failure to behave in ways that the code of conduct considers *desirable* – i.e. the *goal rules* – puts the researchers themselves closer to the centre. It is their *personal* integrity that is affected. The *Netherlands Code of Conduct for Scientific Practice* prohibits inappropriate conduct towards colleagues. But does every harsh, discourteous, or hyper-critical remark made to a colleague constitute a breach of scientific integrity? Here, the scientific forum and scientific integrity clash. The world of science has to allow scientific controversies to be fought out in the scientific forum through argument, claims and proofs, through assertions and contradictions, through replies and rejoinders. The ancient principle that both sides of an argument must be heard also applies in the world of science. Discourteous conduct is irritating and annoying, but it should not be prohibited. Otherwise, a subjective interpretation of integrity – encouraged to some extent by the semantic ambiguity of the term “integrity” – threatens to displace *objective* standards, namely the standards that underlie scientific research, for example honesty, openness, verifiability, and fairness in acknowledging the source of knowledge that one has borrowed from others. For there to be confidence in science, the objective interpretation is more important than the subjective one, which has only arisen owing to the codification of desirable standards of behaviour.

This broad spectrum of scientific misconduct should be the subject of regular discussion within research groups. A public meeting organised in January 2012 by PhD candidates at the Leiden University Medical Centre generated a whole list of different types of misconduct, with researchers being asked to give examples of misconduct and examples of the opposite behaviour. The Committee believes this is one of the simplest ways of ensuring that people remain aware of scientific standards and contraventions of those standards in the research context. Everyday research practice turns out to be the best teacher (see also Section 3.6 and Section 4).

3.4 Is there a relationship between research data management and scientific misconduct?

We know very little, if anything, about the frequency of scientific fraud and the less serious types of scientific misconduct. Only a very limited amount of research has been carried out on this phenomenon. Estimates vary from "absolutely no problem" – as asserted by one of the editors of *Science*, Daniel E. Koshland, who claimed that 99.999% of all articles were accurate and truthful (Koshland, 1987: 141, quoted in Marks, 2009: 182) – to Broad and Wade's claim that for every case of research fraud brought to light, there are 100,000 undiscovered contraventions, both major and minor (Broad and Wade, 1983: 87). They reported concern that the known cases are merely the "tip of an iceberg" (Broad and Wade, 1983: 84). But their claim appears to have been just as exaggerated as Koshland's. Despite the uneasiness that their book caused, and the positive reception it received, Broad and Wade could ultimately identify only 34 specific cases across several centuries (and including some renowned figures such as Newton and Mendel).

A study in 1995 by the *National Institutes of Health* (NIH) in the United States estimated that 1 in 2000 of the research projects financed by the NIK involved scientific misconduct (KNAW, 1995). More recent studies – for example as specified in the ESF's report on *Research Integrity in Europe* (2010) – have led to estimates of one in a hundred researchers being guilty of misconduct over a three to five-year period (Jump, 2010: 20). Analysis using the latest digital methods for detecting plagiarism now indicates a significantly higher percentage of perpetrators than previously assumed, but it remains difficult to interpret these results as long as there is no precise definition of misconduct and plagiarism. The ESF report states, however, that most researchers regard "flagrant cases of deliberate dishonesty" as "rare events" (Jump, 2010: 20). With estimates and claims veering wildly from one extreme to the other, we can only conclude that we simply do not know how big the problem of scientific misconduct actually is. .

The estimates and claims are no less extreme in the Netherlands. Borst considered the likelihood of scientific fraud to be very small and the many thousands and thousands of research projects to be reliable. Van Kolfschooten arrived at somewhat higher figures (mainly for plagiarism) but still believed that "real" fraud was fairly rare (Borst, 1999: 185; Van Kolfschooten, 1996: 14–15). Much depends on how we define research fraud. Do we mean only the most serious cases of FFP, or are we also referring to minor instances of improper behaviour in everyday research practice? Those who say the incidence of fraud is negligible are thinking of rare cases of falsification; those who claim that fraud is widespread are thinking of everyday behaviour, referred to by Köbben as "minor peccadilloes"(Köbben, 2003: 64–79). Is it possible to make a clear distinction between the two types of behaviour: misconduct on the one hand, and everyday sloppiness and minor manipulation on the other? As long as there is not proper, evidence-based research on fraud, all claims are mere speculation. Like every investigation of prohibited and concealed behaviour that only comes to light when the mischief is already done, the necessary research would not be straightforward. Research on scientific fraud is comparable to – and has the same aim – as research on unlawful euthanasia (Van der Maasen et al., 1991), on sexual abuse in the Roman-Catholic church (Deetman et al., 2011), or on fraud in the financial world. The reporting problem – not just hearsay but proper verification (see below) must be solved. So does the "nominator/denominator problem": the ratio of reported and uncovered cases (the nominator) to the total number (the denominator), i.e. including all cases in which something has *not* gone wrong?

Compare the 100 or so reported breaches of scientific integrity investigated by the Dutch universities between 2003 and 2011 – only some of which were found to have actually taken place – to all the research conducted by many thousands of researchers over many years. That ratio might lead us to conclude that scientific misconduct is not a serious problem. If politicians or scientists wish to know the precise nature, extent, and prevalence of scientific misconduct, then the Committee believes that only a thorough scientific study can provide the answer. The study should combine a number of ingenious research methods in order to study a sample of all researchers in all disciplines that is large enough to produce valid conclusions. It would not be sufficient to conduct a simple *self-report* study with general questions about the conduct of *other researchers* (i.e. hearsay) or about *oneself*. Paradoxically enough, the answer to whether or not our confidence in science is misplaced must be found in science itself.

Similar studies have already been carried out in other countries, specifically in the English-speaking world (Fanelli, 2009), but there is no comparable information available for the Netherlands. The Committee therefore believes that serious consideration should be given to carrying out a large-scale study in this country of the prevalence of unacceptable or less desirable practices in research. Such a study might also investigate what role the "pressure to publish" plays: to what extent does that pressure lead not only to a large number of publications (dividing research results into "least publishable units", also referred to as "salami publication") but also to *questionable* practices. In the Committee's view, this recommendation applies primarily to the scientific organisations (TNO, VSNU, NWO, the Academy), which should consider whether and how such a study can be initiated.

3.5 Scientific misconduct and the system of scientific practice

Rather than referring to the "tip of the iceberg" when nobody knows just how large or small the chunk of iceberg below the surface may be, it makes more sense to look at the relationship between incidents and structure. In theory, the Diederik Stapel case is an incident. If we compare it *in detail* with other and similar incidents, however – a process made much easier by knowing the history of scientific fraud – we will learn something about the structure of everyday scientific practice. Rather than speculating about the personal psychological characteristics of perpetrators of scientific fraud, it would be better to investigate their relationship to other researchers and why so many cases of fraud that were ultimately uncovered had long escaped the notice of close colleagues, deans, or journal editors. We must not avoid asking whether – and if so how – our present system of scientific practice influences whether a researcher gives way to the temptation to engage in fraud and deception. Does the growing pressure to publish lead to misconduct, for example, as was suggested after the Stapel case?

What factors in the world of science played a crucial role in the cases of fraud that we know about, and in their only being discovered late on (or too late)? What is the role of the power and reward structures in science, and of the research hierarchy? Has the pressure to publish really increased and what is the precise nature of that pressure? Is it fear of dismissal, of promotions being postponed or cancelled altogether, of having less time available for research, or of a cut in research funding? Is that pressure in evidence at Dutch universities nowadays? Do appointments or promotions depend solely on *number* of publications? Does a linear ranking of the quality of scientific achievement undermine the assessment of that quality? Why does other work – for example close reading of and commenting on colleagues' draft articles and manuscripts – count for less? When is scientific ambition healthy – and when it is not? Would more emphasis on the innovativeness of scientific articles and less emphasis on the number to be produced each year impact the remuneration structure in the world of science? We cannot answer these questions without conducting a specific study, although we can already draw the normative conclusion that, even if work pressure and the "pressure to publish" is higher, there is still no reason or excuse for perpetrating fraud.

Focusing on the system of scientific practice is precisely what anthropologist Jonathan Marks (2009) has done. He considered the phenomenon of scientific fraud from an anthropological perspective, describing specifically the less frequently noted features of day-to-day scientific practice, including "the grapevine", improper exercise of power, and a tendency *not* to believe whistleblowers (or to do so reluctantly). A detailed study of famous and notorious cases of fraud can produce "lessons learned" for the entire academic community and can also generate insights into responsible research data management.

An example will help clarify this. One striking feature of notorious cases of fraud is that the perpetrator is almost always a "lone wolf" who has concealed his research from the eyes and criticism of colleagues. As described in the previous section, the element of peer pressure was missing in such cases. Science is *teamwork*. Group fraud is to all intents and purposes inconceivable and virtually ruled out. Data management can therefore be improved by avoiding situations in which a researcher embarks on a lonely adventure, with too many things left to his discretion. The risk of a researcher "fiddling things" can be minimised by smart strategies in organising research practice and by firm arrangements concerning supervision. That is not to deny brilliant "lone wolves" the opportunity to make discoveries, momentous or otherwise; the point is to make the necessary control mechanisms – for example frequent discussions with other researchers – available even for them.

Studying cases of research fraud in-depth, as the Levelt Committee did in the Netherlands in the Diederik Stapel case, will also reveal weaknesses in journal editorial practices, for example the problem that peer reviewers sometimes spend too little time reading and reviewing submissions or pass these important duties on to their assistants (Levelt et al., 2011, 2012; Marks, 2009: 181). In some cases, journal editors do not stick to their own rule that a submission will be published only if there is a full set of research data available for verification purposes; instead, they restrict themselves to assessing the texts submitted. This also applies to the work carried out by doctorate committees when evaluating dissertations. The Committee believes that studying the lessons that can be learned from fraud – based on both familiar and less familiar cases – and from normal practices in scientific research will be more valuable than simply identifying fraud when it has already taken place (however one defines it). Confidence in science does not depend on identifying a number of rotten apples; more relevant is the *basket* holding *all* the apples.

Based on his study of familiar and less familiar cases of fraud, Marks identifies four "lessons learned" (Marks, 2009: 188- 190). The first is that someone can easily get away with fraud if the procedure for identifying and tackling it is ineffective, if the perpetrator himself continues to take legal action (as in the Baltimore case), and if the authorities have an interest in hushing up the fraud. The second lesson is that it is not only the perpetrator's research career that suffers, but also the whistleblower's; close colleagues or fellow researchers are therefore not much inclined to bring up the matter. The third lesson is that researchers are all too willing to give colleagues the easy way out by denying that it is a case of "real" fraud, but merely "mistakes" or "sloppiness" in the research. In many instances, the co-authors themselves makes these excuses to avoid sharing the blame. Marks judges such defences harshly because they make scientific work unreliable: "After all, once you have established that your colleague's work is not reliable, it really doesn't matter why. If some scientists don't do good research, it is difficult to maintain that they should nevertheless still be employed and receiving grants, much less that you want to continue collaborating with them. The problem with the 'incompetence defence' then, is that it implicitly raises a question about the rest of their work…. To say someone is a sloppy researcher whose work is riddled with mistakes is not a compliment, and it immediately raises the questions of why you are associated with such a person, how competent the rest of their research has been and why they should remain at work. I can think of no other profession in which that would be tolerated" (Marks, 2009: 189). This lesson refers back to the issue of the personal responsibility of co-authors of fraudulent or questionable articles – an issue that hardly ever comes up in the current discussion of scientific fraud.

The fourth and final lesson noted by Marks concerns the hierarchy within the scientific world. It is precisely when the fraud is perpetrated by highly regarded researchers who exercise a great deal of power within their field that everything possible is done – at least at first – to deny that fraud has in fact taken place or to launch a counterattack against the whistleblowers (the Burt case is a good example). The burden of proof is a heavy one because people trust almost blindly in the good name of the research concerned. In a number of cases, that trust has been entirely unjustified, showing us the need for effective channels to report suspicions of fraud. Tilburg University learned that lesson after the Diederik Stapel case, in which PhD candidates were apprehensive about reporting their suspicions of their own dean. Marks' conclusion concerning research assessment is therefore short and sweet: "The only relevant criteria for establishing the quality of scientific work is whether it is competent and honest; all else is polemics" (ibid. 190).

Whether and how we can combat scientific fraud is not necessarily a question of the perpetrator's personal characteristics, but equally – at least where the world of science is concerned – a question of the structure and system of everyday scientific practice and, in that context, primarily a question of honest and responsible research data management. The Committee believes that improving how researchers handle their data, as recommended in the previous section, will make scientific research less susceptible to fraud. Poor and sloppy scientific practice, particularly in the initial phase of the research, is not the same thing as fraudulent science, nor is there a direct causal relationship with serious kinds of fraud. However, by combating poor and sloppy research practice and by promoting effective data management and proper intellectual control of research data will reduce the likelihood of serious contraventions.

The meaning of scientific work

Amidst all the uproar about scientific integrity and reliability, it is worth remembering that the aim is always to replace current scientific findings and insights by new and better ones. Max Weber already pointed this out in 1919 when he wrote: "That is the fate to which science is subjected; it is the very meaning of scientific work." (Weber, 1919). Science builds on earlier science. Today's findings and insights are processed and transformed into those of tomorrow, and today's knowledge will often become virtually unrecognisable in the distant future. Science does not need to be perfect in order to progress.

One can find fault with every study and every data collection. Data collection, analysis, and reporting are always coloured to a certain extent by the researcher's initial hypothesis. This explains why even giants in the history of science are known to have omitted data that did not fit in with their theories in order to gain a clearer picture (Klotz, 1992; Broad and Wade, 1983). It was still possible to build on their findings, however, because their theories turned out to be effective ones. In this way, science is ultimately "self-refining".

The balance is a delicate one, of course. Where it lies – how much freedom of interpretation is possible, i.e. when should we consider an experiment to have "failed" and to be too unimportant to report on – depends on the scientific field and the relevant objectives. In research that explores the boundaries of our understanding, we can allow for a great deal of speculation – everybody working within the discipline knows this and is aware of the tentative nature of the research. The most important thing in such cases is creativity. At the other extreme, when a therapeutic medical experiment has consequences for the registration of a medicinal product and therefore a potential impact on millions of patients, we must adhere to a strict, pre-determined protocol must without any deviation; here, scientific imagination is entirely unacceptable.

In the view of the Committee, informal peer pressure offers the best guarantee of our dealing properly with this delicate balance, followed by formal peer review. In other words, researchers need to discuss their work with scientific colleagues before, during, and after the research; those colleagues will recognise all too fanciful explanations and analyses and can clarify to the researchers involved that their explanations and analyses are either too far removed from the data or that they are indeed interesting but should best be published as explicitly hypothetical.

3.6 Recommendations

The recommendations set out in Section 2 of this Advisory Report on responsible data management are also relevant to the problem of scientific misconduct. Responsible research conduct and adequate supervision and verification are, after all, the first step towards promoting scientific integrity and preventing scientific fraud. The precautionary principle also applies to scientific integrity, namely the need to do everything necessary to reaffirm confidence in science. This requires the universities to uphold a proper "research culture". The precautionary principle obliges the universities to create the right conditions for promoting research integrity, for example by making manpower and funds available for proper data storage and archiving.

Furthermore, the basic rules of scientific practice are clear and do not need to be amended. The Committee does not consider it necessary for them to be tightened up or for new legal measures to be imposed alongside these rules. The Committee also believes that satisfactory fraud prevention measures are already provided for in existing statutory rules and in the obligations that employment contracts impose concerning conduct befitting a good employee. It does not consider that introducing new regulatory agencies or government-appointed bodies to monitor scientific integrity will contribute to solving the problem. After all, we have yet to actually define the problem properly and its nature and scope will remain unclear until the topic of scientific fraud has been properly investigated. Caution is also advised with respect to new local monitoring measures in the form of "quality protocols" at universities and university medical centres; these are only useful if they solve a clearly defined problem.

The Committee therefore urges the research organisations VSNU, NWO, TNO and the Academy to conduct or commission an investigation into scientific misconduct in the Netherlands in order to ascertain the nature of such misconduct and its prevalence. Such clarity is needed as a defensive measure in order to remove any doubt concerning the integrity of research.

3.6.1 Awareness and familiarity with procedures in cases of scientific misconduct

What must certainly be improved, however, are familiarity with the rules among the broad group of researchers and collective awareness of the importance of responsible research conduct and scientific integrity in boosting confidence in science. Awareness begins by doing more to publicise the rules and by clarifying their purpose. The Dutch universities, working within the context of the VSNU, seized on the Diederik Stapel case in order to carry out a self-analysis. Even if we believe that lies and deception will never be entirely eradicated, it may still be useful – and indeed advisable in view of that belief – to create underlying conditions that make fraud less tempting and increase the chance of its being detected.

The self-analysis showed that some universities had not developed proper procedures for dealing with cases of minor or serious misconduct. Where procedures had been put in place, it was sometimes unclear how and where suspected cases should be reported. The necessary procedures have taken shape over the past few years, however, an outcome of the complaints procedures for "A" and "B" government organisations, as currently provided for in Book 9 of the General Administrative Law Act [*Algemene Wet Bestuursrecht*]. There is therefore already a statutory basis for (mandatory) handling complaints regarding scientific integrity. The "special" [*bijzondere*] universities have adopted this system, or they can choose to develop their own regulatory system under private law. The procedures gave rise to various problems, for example the fact that the responsible administrators (deans, vice-chancellors, executive boards) are simultaneously responsible for independently investigating cases of fraud, but the VSNU succeeded in rectifying these in a National Model Complaints Procedure (2012) [*Landelijk Model Klachtenregeling 2012*] that applies to all the universities.

3.6.2 Familiarity with the Netherlands Code of Conduct for Scientific Practice

The 2004 version of the *Netherlands Code of Conduct for Scientific Practice* [*Nederlandse Code Wetenschapsbeoefening*] has also been revised at the VSNU's initiative with a view to making it clearer and more effective. It is important that the *Code of Conduct* not only spells out the obvious prohibitions (as in the notorious cases of FFP) but also, and more specifically, makes choices concerning the larger "twilight area" of minor breaches of scientific integrity. Researchers need clarity about how serious certain conduct is within the context of scientific integrity, and what the appropriate sanctions are. They need to know how the co-authors of fraudulent articles will be assessed. Can they be considered to be complicit in the fraud? There are few rules pertaining to any statute of limitations for scientific fraud: in cases of serious fraud, such as the Diederik Stapel case, it is worth checking up on articles even if they were published more than ten years ago; for other articles, a ten-year statute of limitations would seem practical and fair.

It is important to draw the attention of the entire research community to the *Netherlands Code of Conduct for Scientific Practice*, and to make scientific values and standards available to new appointees in an easy-to-use format, both in English and in Dutch, as Erasmus University Rotterdam does. It would be imprudent to expect this to be enough, however. As the Dutch proverb says, "paper is patient" and practice is more important than doctrine. The Committee therefore advises making scientific integrity an enduring topic of discussion in academic programmes (see Section 4), in the research community itself, and among the managers and directors of research organisations.

3.6.3 Research institutes

Research institutes are basically responsible for paying sufficient attention to integrity in research practice. One way that they can do this is by raising the topic during performance appraisal interviews with researchers. They can also, in their own way, initiate an open discussion about research integrity, or aspects thereof, for example during research conferences and seminars (see Section 4). They can raise awareness by combining discussions of research skills, current research practices, and integrity issues. The regularity and frequency of such discussions should be compatible with the particular field of research and existing discussion practices.

3.6.4 Universities

The *Standard Evaluation Protocol* (SEP), which is used during external reviews, should offer scope for assessing the interest that universities and research institutes take in scientific integrity and how they keep that interest alive.

One specific way in which universities can emphasise the importance of scientific integrity within their institution is to introduce a professional oath (or solemn affirmation) whereby a researcher officially declares that he will comply with the *Netherlands Code of Conduct for Scientific Practice* in all research that he carries out, whether now or in the future. Although such an obligation is already inherent in the employment contracts between researchers and their employers – so that an additional declaration is not strictly necessary – it becomes more explicit when they are asked to sign an official declaration, raising their awareness of the standards and values that underpin scientific practice. One objection to a professional oath or affirmation is that it is not really necessary for honest researchers and that it will not prevent those with ill intentions from behaving dishonestly. On the other hand, like doctors, lawyers and other professionals, researchers who have taken such an oath or affirmation can fall back on it if it later becomes desirable or necessary to discuss integrity problems with their supervisors or research coordinators. A professional oath or affirmation is not merely of symbolic value, in other words.

Since discussion of a mandatory oath or affirmation has scarcely begun, the Committee recommends that each university develop its own policy in this regard, which, in a few years' time, can be combined into a joint approach. At the moment, the Committee believes it is more important to offer different formats for an official declaration of this kind (oath or affirmation; at the start of a researcher's career or at its conclusion when receiving his or her doctorate; with or without a compulsory course, see Section 4) than to impose an obligation on all organisations involved in research to introduce a professional oath.

3.6.5 The role of the Academy

As indicated above, the Committee urges the Academy and other research organisations to conduct (or commission) a study of scientific misconduct in the Netherlands in order to clarify the nature of such misconduct and how prevalent it is. Such clarity is needed as a defensive measure in order to remove any doubt concerning the integrity of research.

The Academy can also play an important role in disseminating information on dilemmas involved in scientific integrity and other related issues. In 2000, the Academy published the booklet *Scientific Research: Dilemmas and Temptations* [*Wetenschappelijk onderzoek: dilemma's en verleidingen*]. The booklet was so well received that a revised second edition was published in 2005. The final section, "Prevention and Remedies", offers a number of valuable recommendations on training, quality assessment and quality evaluation, the appointment of confidential counsellors and integrity committees, and scientific practice in general. It is now 2012, and many of those recommendations have still not been complied with; they could in fact be repeated almost verbatim in the present Advisory Report. The bodies responsible did not take a keen interest in the topic In 2005, although there was considerable demand for copies of the booklet. It is now out of print and the Academy does not intend reprinting it in its 2005 form. The Committee nevertheless considers it important for a similar booklet to be published, updated and revised, with new cases and new analyses. The new publication could include the revised 2012 version of the *Netherlands Code of Conduct for Scientific Practice*. It is up to the Academy (whether or not in collaboration with TNO, VSNU, and NWO) to produce such a booklet and to ensure that it is distributed widely. The booklet *On Being a Scientist* can also be included (see Section 4).

The most important recommendation that the Committee can make with respect to scientific integrity is to ensure that it becomes a topic of enduring interest throughout the world of research, in particular within the responsible bodies. The Committee advises the Board of the Academy to investigate, in four years' time, the extent to which progress has been made in research data management and in the problems inherent in scientific integrity, and whether interest in these matters is perhaps waning, as it did six years ago.

3.6.6 The researchers

Last but not least, the researchers themselves have a duty to conduct themselves with integrity in their work. That duty follows from their personal decision to pursue science as a profession and to conduct themselves in accordance with its standards. Above all, they must internalise those standards so that integrity goes without saying and becomes second nature to them in their research practice. Such internalisation is only possible if the subject of integrity is discussed repeatedly in the workplace (see Section 4). Integrity is therefore vital to fostering the trust that researchers have in one another and the trust that society should rightly be able to place in the results of research.

4. PREPARATION FOR SCIENTIFIC RESEARCH IN UNDERGRADUATE AND POSTGRADUATE PROGRAMMES AND DURING RESEARCHER TRAINING

4.1 Introduction

Because they serve as the environments in which new generations of researchers are trained, universities have the task of inculcating not only technical skills but also of teaching students to be good researchers in the ethical sense – in other words, researchers who are responsible, competent, and act with integrity. This gives universities a twofold task that combines, in a manifestly recognisable form, the two main topics of this Advisory Report: "responsible research data management" and "scientific integrity". The Committee believes that these two essential aspects of research should not be omitted from academic programmes or from researcher training and can best be dealt with in close combination. This generally used to take place within the master-apprentice system, with both technical research skills and a scientific mindset and values being passed on almost automatically. It has become more difficult to do this now, with the nature of university education and research having undergone dramatic changes and with many more people involved. The one-to-one model is only effective in small-scale educational and research settings; elsewhere, new methods are being introduced. We make a number of suggestions in that regard in this section. The Committee assumes that both aspects of research (responsible conduct and integrity) deserve attention, preferably at the same time, but acknowledges their design and transmission depend so much on the nature of the discipline and the setting in which teaching and research take place (large-scale or small-scale; a great deal of prior knowledge and training or only a little; the nature of the research that is to be carried out) that it is better to leave the question of "how" to the researcher training programmes themselves.

4.2 Undergraduate and postgraduate programmes

Educational programmes in the Netherlands include training in "research skills". Undergraduate degree programmes pay little attention to the topic of "integrity in science and in scientific research". In the past decade, however, interest has grown in the position of the student as a participant in university education; this can be seen, for example, in the recent *Code of Conduct* for students, which forms the basis for the Dutch higher education system's "education contracts" (in which students undertake to do their best to obtain a certain number of course credits each academic year – in other words: they assume a "best-efforts obligation").

Students taking a Bachelor's degree in many fields of study are obliged to carry out research projects, with the emphasis being on acquiring research skills, methods, and techniques. This is probably too brief an introduction to research – certainly following the restructuring of university education in 2002 – to also cover issues of scientific integrity. In virtually all fields of study, however, the first trimester of the Bachelor's degree programme includes a unit on "study and publication skills" (how to use the library, source references, digital data acquisition, etc.). These courses teach students from the very beginning that their own assignments, papers, etc. must always provide source references if material is taken from books or articles. They are made to understand that plagiarism is entirely unacceptable. This message is meant to counterbalance the growing number of ways students can download material from the Internet and the resulting practice of utilising that material without citing the source. Software that can detect plagiarism is probably more effective in this context than a lecture on what is and is not permitted in terms of source referencing. The growing, unacceptable habit of utilising Internet material without indicating the source must be detected and the relevant students broken of this habit at the earliest possible stage of their programme.

The Committee believes that Bachelor's programmes should stress the intrinsic importance of responsible research conduct and scientific integrity with fledgling researchers at every available opportunity:

- library courses: dealing with sources and citations: what is plagiarism?
- Bachelor's and Master's degree research: data collection (responsible conduct and truthfulness, neither fabrication nor falsification);
- methods and statistics: devoting explicit attention to pitfalls and dilemmas (missing observations, outliers) and to "how to lie with statistics".

Master's degree programmes should provide instruction in ethics. This should involve more than teaching students techniques; it should include acquainting them with legislation, regulations, and codes of conduct on the one hand (depending on the specific profession for which the student is being trained), and cultural transmission on the other. The latter requires using other educational methods, such as reflection and debate. The point is for the student to acquire a specific scientific mind-set. It is only on this basis that a programme of researcher training can be established.

4.3 Researcher training

Only a small percentage of students choose a career in research (an estimated five percent). We need to cherish these students because they have chosen a profession that makes heavy demands on their curiosity, creativity, truthfulness, honesty, openness, critical thinking, ability to work in teams, and other traits typical of someone with a scientific mind-set. It is all the more striking, then, that the preparation and supervision of these young researchers fails in a number of respects. Every researcher in the Netherlands who is just embarking on his/her career should read the booklet *On Being a Scientist: Responsible Conduct in Research* (National Academy of Sciences, 2005) and discuss it in a group; this would serve as both an incentive and a guide. The universities could provide all new researchers with a copy. (See also the previous section on the task of the Academy.)

At the moment, few if any doctoral programmes commence with an introductory course focusing specifically on how researchers should manage research data (including statistical skills) and also addressing aspects of scientific integrity. This is regrettable, especially because a considerable percentage of these young researchers come from abroad and have often not taken their Bachelor's or Master's degree in the Netherlands. Delft University of Technology recently started a research course for newly appointed researchers. This could serve as an example for other universities and/or for large research schools and university medical centres that employ large numbers of PhD candidates.

The Committee believes that proposals for courses of this kind are useful, provided that they preserve the link between integrity in research and research practice. The format of such a course will differ from one discipline and research situation to another. Courses naturally require a certain number of participants before they can go ahead, but in small research groups it might be difficult to reach that level of enrolment. In such cases, on-the-job training is probably more appropriate. Also, researchers will be more interested in how to deal with the practical problems of their work when they have actually come to grips with such problems, rather than beforehand, at the beginning of the research period. In the view of the Committee, the best approach would therefore be a course that deals with these topics in a manner close to the actual practice of research and not as a separate component of the researcher training programme. PhD supervisors and other senior officials can transmit the standards and values applying to proper research "naturally", with students learning by imitation. One point to bear in mind is the risk that these matters will once more escape attention, as has happened in the past. The task of transmitting research standards and values should therefore be seen as an integral part of research *practice,* for example involving annual or semi-annual public discussion of research practices with young researchers (see also Section 3). The advantage of such an approach is that discussions of theoretical and practical integrity issues would involve not only PhD candidates but also postdocs and senior researchers. Practice is the best way of learning in this regard. Bureaucratisation of the research period – with students obliged to take separate courses on ethics, scientific integrity, responsible data management, and quality protocols – might be too much of a good thing and would detract from the pleasure involved in the adventure that is research.

4.4 Recommendations

To summarise: the Committee recommends that the ordinary curriculum during Bachelor's and Master's degree programmes should devote attention to and provide scope – if that is not already the case – for dealing with specific questions regarding responsible research conduct and scientific integrity. The Bachelor's degree component should focus primarily on the topics of plagiarism, data collection, and statistics, as described in Section 4.2. The Master's degree component would preferably concentrate on reflection and discussion of the ethical aspects of research (a broader topic than scientific integrity alone).

The Committee believes that researcher training programmes or the initial phase of PhD programmes should cover the topics of responsible research data management (including statistical and other skills) and scientific integrity. Just *what* material they should actually cover during that phase is not something that a single rule or recommendation can indicate; much depends on the scientific domain, the discipline, and the specific research setting. In the Committee's view, the ultimate goal – being a good scientist – is more important than having a prescribed route for achieving that goal in every scientific discipline.

5. SUMMARY

Should we worry about scientific misconduct?

Responsible research conduct and scientific integrity are indispensable to the pursuit of science. Public trust in science depends on them. Should the research community be worried about the integrity of scientific research, in particular the way in which researchers manage their data? Do researchers have a tendency to present their work in a more positive light than is actually justified? How can we encourage researchers to conduct themselves more responsibly and with greater integrity when managing their data? The nature of these questions led the Board of the Royal Netherlands Academy of Arts and Sciences to install an advisory Committee on Scientific Research Data.

Science, in the end, is self-refining

In the Committee's view, researchers have an obligation to counteract any realistic threat to public trust in science as quickly and as vigorously as possible. As a precautionary measure, and with a view to promoting unwavering trust in science, researchers should examine their own work and verify whether it in fact meets all of the stringent requirements that have been set for research. But the pursuit of science does not have to be perfect in every way. The system is organised so that the research results and findings of today are reconfigured and transformed into those of tomorrow. Creativity plays a major role in this process, along with a highly critical attitude. This is how research ultimately becomes self-refining.

Accessibility and vitality

The Committee's main recommendation is that access to research data must be guaranteed, bearing in mind the many different ways that researchers conduct their work. The Committee also recommends developing procedures that make the existing rules on integrity inherent in research practice; it emphasises that this should be a key point of attention in academic training.

Responsible research data management

After reviewing the findings of a survey of researchers working in the field, the Committee concludes that the differences between and within disciplines are too large to arrive at general conclusions about responsible data management in research. Any problems that arise in that context, and any potential solutions, should be examined and brought up for discussion within the relevant discipline.

Checks and balances, before and after

An analysis of the research cycle makes clear, however, that the initial phase of the research process (designing the test environment, data collection, statistics, analysis) is relatively exploratory, with verification being left to the individual researcher or research group. It is vital in this phase to deal responsibly with research data. In a number of disciplines, verification takes place largely an *ex post* basis, primarily by peer review of the final report, and – after publication in professional journals – by the scientific community. Where *ex post* verification is more problematic – perhaps because the data is difficult to access, either because it is rare and unique (the result of many years of archive work), or because only a handful of researchers are working on the same subject – the focus should obviously come to lie more heavily on the initial research phase. Conversely, if the data is easy to access, if a full account is given of the methods used, and if there is a functioning forum active within the discipline, then there is less need for verification in the early phase of research.

Mutual exchanges

It is up to the disciplines to examine how *ex ante* and *ex post* verification are related and to adjust this relationship where necessary. Disciplines can certainly learn from best practices in other disciplines and adopt their control mechanisms, for example by keeping diaries, logbooks or lab journals, accounting for data, conducting research in teams, and applying peer pressure *before* the peer review phase.

Investigating research data management

The Committee recommends making a systematic examination of how different disciplines deal with research data in actual practice. The study could, for example, involve a number of randomly selected research projects in each discipline. By mapping the entire "life cycle" of these projects, the study will not only reveal how researchers in fact manage research data but will also offer hard evidence of crucial differences between disciplines and how each one has solved certain recurring problems in that regard. In addition, the study can consider specific technical aspects of data management within particular disciplines. It is up to the umbrella organisations – the Association of Universities in the Netherlands (VSNU), the Netherlands Organisation for Scientific Research (NWO), the Netherlands Organisation for Applied Scientific Research (TNO) and the Academy (KNAW) – to decide how a systematic study of this kind should be carried out.

The responsibility for research data lies at three different levels: 1) the individual researcher; 2) the research institution; and 3) the informal networks that play such a vital role in the research community.

The individual researcher

Individual responsibility

A well-trained, discernibly scientific mind-set is a necessity for all researchers. Inherent to this mind-set is the insight that they themselves bear primary responsibility for responsible conduct in their research and in data collection, storage, management, processing and reporting.

The research institution

It is up to research institutions to create and maintain a climate in which a mind-set conducive to responsible conduct can thrive. A culture of critical thinking makes all the difference in this respect. Even before papers are peer-reviewed, fellow researchers most closely involved in the project, research leaders, PhD supervisors and so forth should apply the relevant peer *pressure*. Research institutes should support the research process throughout the course of a project, for example by providing assistance with data management and statistical analysis. Faculty chairs and the directors of research institutes can make responsible data management one of the points covered in researchers' annual performance appraisals. Institutes can also organise yearly seminars focusing on the problems involved in data management and data processing. Such events will help maintain awareness of the need for responsible research. The councils and boards of research institutions should ensure that researchers do in fact address the issue of responsible conduct in the primary research process, for example by arranging the necessary funding and staff. Responsible data management should be a permanent item in the Standard Evaluation Protocol (SEP) adhered to by VSNU, NWO and the Academy, which is used to evaluate both research and the institutions themselves.

Control is needed

Informal networks

Local, national and international networks promote frequent communication between researchers. Researchers themselves play many different roles in these networks: they are peers; they assess the work of other researchers in their discipline; they review journal submissions; they edit scientific publications; they organise conferences and other academic gatherings; and they sit on doctoral examining committees and external evaluation committees. The members of the research community therefore bear the huge responsibility of remaining critical of the standards and values applied in their discipline for data management. Professional associations and journals are important nodes in informal networks. Journal editors should ensure

Strong informal networks

that authors can account for their research results. In peer-reviewed journals, the underlying research data should, as a rule, be accessible to other researchers. Rules concerning co-authorships should be reconsidered in each discipline and, where necessary, tightened up so that it is clear what co-authors can be required to know about the research data and methods and where the limits of their responsibility lie.

Scientific integrity

Scientific values

Scientific integrity is a specific standard of behaviour linked to the position of researcher. It means acting in accordance with scientific values, for example truthfulness, honesty and openness in reporting, even when no one is monitoring the researcher's actions. Science rejects certain behaviours that are permitted in other segments of society. In many social institutions it is acceptable practice to copy someone else's work, but in science that is not allowed, at least not unless the source is cited. Science's leading values differ from those of other sectors of society in that respect.

Criteria for scientific misconduct

In what situations can we say that scientific integrity has been violated, and what criteria can we apply when assessing possible violations? There are three such criteria, well known to all researchers: Fabrication, Falsification and Plagiarism, referred to simply as FFP:

- Fabrication: fraud; inventing or faking research data;
- Falsification: manipulating data or presenting data erroneously, for example by neglecting to include negative outcomes;
- Plagiarism: quoting passages verbatim or "stealing" or using someone else's ideas without citing their source.

The size of misconduct is unknown

Very little if anything is known about the frequency of violations against scientific integrity. Only a very limited amount of research has been carried out on this phenomenon. Estimates vary from "never" to claims that for every case of research fraud brought to light, there are approximately 100,000 undiscovered violations, both major and minor. With estimates and claims veering wildly from one extreme to the other, we can only conclude that we simply do not know how big the problem of scientific misconduct actually is. The estimates and claims are no less extreme in the Netherlands. Much depends on how we define research fraud. Do we mean only the most serious cases of FFP, or are we also referring to minor instances of improper behaviour in everyday research practice? Those who say that the incidence of fraud is negligible are thinking of rare cases of falsification; those who claim that fraud is widespread are thinking of everyday behaviour. As long as there is no proper, evidence-based research on fraud, all claims are mere speculation.

Research into the size of misconduct

The Committee therefore urges the research organisations VSNU, NWO, TNO and the Academy to conduct or commission an investigation into scientific misconduct in the Netherlands in order to ascertain the nature of such misconduct and its prevalence. Such clarity is needed as a defensive measure in order to remove any doubt concerning the integrity of research.

Warning against lonely adventures

One way to improve matters is to disabuse researchers of the notion that research is a solo adventure in which much is simply left to their discretion. The risk of a researcher "fiddling things" can be minimised by smart strategies in organising research practice and by firm agreements concerning supervision. That is not to deny brilliant "lone wolves" the opportunity to make discoveries, momentous or otherwise; the point is to make the necessary control mechanisms – for example frequent discussions with other researchers – available even for them.

Scientific magazines have a responsibility as well

Studying cases of research fraud in-depth will also reveal weaknesses in journal editorial practices, for example the problem that peer reviewers sometimes spend too little time reading and reviewing submissions. That also applies for the work of doctoral examining committees in evaluating dissertations.

No new rules or controlling bodies

The Committee does not see a need to tighten up existing rules of conduct on the point of research integrity. The rules are already clear and straightforward enough. The Committee also believes that satisfactory fraud prevention measures are already provided for in existing statutory rules and in the obligations that employment contracts impose concerning conduct befitting a good employee. It does not consider that installing new regulatory agencies or government-appointed bodies to monitor scientific integrity will contribute to solving the problem. After all, we have yet to actually define the problem properly and its nature and scope will remain unclear until the topic of scientific fraud has been investigated. Caution is also advised with respect to new local monitoring measures in the form of quality protocols at universities and university medical centres. These are only useful if they solve a clearly defined problem.

Awareness

What is more important is to ask how researchers can be mindful of the rules and comply with them in their everyday research work. In real-life situations, borderline cases do occur and the grey area is fairly large. Researchers across the board need to be more familiar with the rules. This will help them understand the importance of conducting themselves responsibly and with integrity. Researchers will be more mindful if we do more to familiarise them with the rules, clarify their purpose, and enforce them properly.

Even if we believe that lies and deception will never be entirely eradicated, it may still be useful – and indeed advisable in view of that belief – to create underlying conditions that make fraud less tempting and increase the chance of its being detected.

Levels of responsibility

Emphasising integrity as an individual trait draws attention away from the structural component that is also part and parcel of integrity. Like the question of responsible data management, integrity involves striking the right balance between individual responsibility and the responsibility of organisations for the circumstances in which individuals work. Here again, a distinction can be made between different levels of responsibility.

The researchers

Internalising the standards

The first level is that of the researcher. Researchers have an obligation to conduct themselves with integrity in their work. That obligation follows from their personal decision to pursue science as a profession and to conduct themselves in accordance with its standards. Above all, they must internalise those standards so that integrity goes without saying and becomes second nature to them in their research practice. Internalisation is only possible if the issue of integrity is discussed repeatedly in the workplace. Integrity is therefore vital to fostering the trust that researchers have in one another and the trust that society should rightly be able to place in the results of research.

The research institution

Attention for the code

It is certainly important to repeatedly draw the attention of the entire research community to the *Netherlands Code of Conduct for Scientific Practice*, and to make scientific values and standards available to new appointees in an easy-to-use format, both in English and in Dutch. It would be imprudent to expect this to be enough, however.

The Committee therefore advises making scientific integrity an enduring topic of discussion in academic programmes, in the research community itself, and among the managers and directors of research organisations.

Talk about it

Research institutes are basically responsible for paying sufficient attention to integrity in research practice. One way that they can do this is by raising the topic during performance appraisal interviews with researchers. They can also initiate a discussion about research integrity, or aspects thereof, for example during research conferences and seminars.

Universities can emphasise the importance of scientific integrity in their institution by introducing a professional oath whereby the researcher undertakes to comply with the *Netherlands Code of Conduct for Scientific Practice*. Although such an obligation is already inherent in the employment contracts between researchers and their employers, it becomes more explicit when they are asked to sign an official statement, raising their awareness of the values and standards that underpin the pursuit of science. The Committee recommends universities to develop separate policies on this score that will eventually converge into a consistent policy line. At the moment, the Committee believes it is more important to offer different formats for an official statement of this kind (oath or affirmation; at the start of a researcher's career or at its conclusion when receiving his or her doctorate; with or without a compulsory course) than to impose an obligation on all organisations involved in research to introduce a professional oath.

Professional oath or promise

Because they serve as the environments in which new generations of researchers are trained, universities have the task of inculcating not only technical skills but also of teaching students to be good researchers in the ethical sense – in other words, researchers who conduct themselves responsibly, competently, and with integrity. This gives universities a twofold task that combines, in a manifestly recognisable form, the two main topics of this report: "responsible data management" and "scientific integrity". The Committee believes that Bachelor's programmes should stress the intrinsic importance of responsible conduct and integrity with fledgling researchers at every available opportunity:

To kindle at an early age

- library courses: dealing with sources and citations (what is plagiarism?);
- Bachelor's and Master's degree research: data collection (responsible conduct and truthfulness, neither fabrication nor falsification);
- methods and statistics: devoting explicit attention to pitfalls and dilemmas (missing observations, outliers) and to "how to lie with statistics".

Master's degree programmes should provide instruction in ethics. This should involve more than teaching students techniques; it should include acquainting them with legislation, regulations, and codes of conduct on the one hand (depending on the specific profession for which the student is being trained), and cultural transmission on the other. The latter requires using other educational methods, such as reflection and debate.

Only a small percentage of students choose a career in research. They have chosen a profession that makes heavy demands on their curiosity, creativity, truthfulness, honesty, openness, critical thinking, ability to work in teams, and other traits typical of someone with a scientific mind-set. At the

Researcher training

moment, few if any doctoral programmes commence with an introductory course focusing specifically on how researchers should manage research data (including statistical skills) and also addressing aspects of scientific integrity. The Committee believes that proposals for courses of this kind are useful, provided that they preserve the link between integrity in research and research practice. The format of such a course will differ from one discipline and research situation to another.

BIBLIOGRAPHY

Altman, D.G., “The scandal of poor medical research” in: *British Medical Journal,* 1994: 308:283

Babbage, C., *Reflections on the Decline of Science in England, and on Some of Its Causes,* London: B. Fellowes, 1830; in: C. Babbage, *Science and Reform: Selected Works,* ed. A. Hyman, Cambridge/New York: Cambridge University Press, 1989, pp. 122 ff

Bauer, H.H., “The knowledge filter” in: H.H. Bauer, *Scientific Literacy and the Myth of Scientific Method,* Urbana: University of Illinois Press, 1992

Borst, P., *De vioolspelende koe en andere muizenissen,* Amsterdam: Bert Bakker, 1999

Borst, P., “De baas wist van niets”, in: *NRC Handelsblad,* 16 November 2002

Broad, W. and N. Wade, *Betrayers of the Truth, Fraud and Deceit in the Halls of Science,* New York: Simon and Schuster, 1983

Brown, T., *Imperfect Oracle, the Epistemic and Moral Authority of Science,*

Pennsylvania: The Pennsylvania State University Press, 2009

Callahan, D., *What Price better Health? Hazards of the Research Imperative,* California: Milbank Books on Health and the Public, University of California Press, 2003

Carter, S.L., *Integrity,* New York: Basic Books, 1996

Chiong Meza, C., *De Nederlandse Universiteiten 2012: Feiten en Cijfers 6,* The Hague: Rathenau Institute, 2012

Deetman, W. et al., *Seksueel Misbruik van Minderjarigen in de Katholieke Kerk,* Amsterdam: Balans, 2011

Durkheim, E., *De la Division du Travail Social,* Paris: Presses Universitaires de France, 1893; (10th ed., 1978)

ESF/ALLEA, *European Code of Conduct for Research Integrity,* Amsterdam/Strasbourg: European Science Foundation/All European Academies, 2011

Fanelli, D., “How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data”, Plos ONE, (4/5), e 5738

Goodstein, D., *On Fact and Fraud, Cautionary Tales from the Frontlines of Science*, Princeton: Princeton University Press, 2010

Groot, A.D. de, *Methodologie, Grondslagen van Onderzoek en Denken in de Gedragswetenschappen,* The Hague: Mouton and Co 1961; 1981; English ed. 1969

Groot, A.D. de, *Academie en Forum, over Hoger Onderwijs en Wetenschap,* Amsterdam: Boom, 1982

Haack, S., *Defending Science within Reason, between Scientism and Cynicism,* New York: Prometheus Books, 2003

Heilbron, J., M. van Bottenburg and I. Geesink, *Wetenschappelijk Onderzoek: Dilemma’s en Verleidingen,* Amsterdam: KNAW, 2000

Heilbron, J., *Wetenschappelijk Onderzoek: Dilemma’s en Verleidingen,* Amsterdam: KNAW 2005, tweede druk http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20041076.pdf

Huberts, L.W.J.C., *Integriteit en Integritisme in Bestuur en Samenleving,* inaugural lecture, Amsterdam: Vrije Universiteit, 2005

Jump, P., “The good Conduct Guide” in: *Times Higher Education Supplement,* August 5th 2010 http://tinyurl.com/37vywcs

Kevles, D.J., *The Baltimore Case, a Trial of Politics, Science and Character,* New York: W.W. Norton, 1998

Kirkels, V. (ed.), *Eed van Hippocrates, nog van deze Tijd?* Nijmegen: Valkhof Pers, 2004

Klotz, I.M., *Cooking and trimming by scientific giants* , Faseb Journal, Vol. 6 pp 2271 – 2273, 1992

KNAW, NWO, VSNU, *Notitie inzake Wetenschappelijk Wangedrag,* Amsterdam: KNAW, 1995

KNAW, NWO, VSNU, *Notitie Wetenschappelijke Integriteit, over Normen van Wetenschappelijk Onderzoek en een Landelijk Orgaan voor Wetenschappelijke Integriteit,* Amsterdam: KNAW 2003

KNAW, *Wetenschap op Bestelling, over de Omgang tussen Wetenschappelijk Onderzoekers en hun Opdrachtgevers,* Amsterdam: KNAW, 2005

Köbben, A.J.F., *Het gevecht met de Engel, over Verheffende en minder Verheffende Aspecten van het Wetenschapsbedrijf,* Amsterdam: Mets en Schilt, 2003

Kolfschooten, F. van, *Valse Vooruitgang, Bedrog in de Nederlandse Wetenschap,* Amsterdam: Pandora/Contact, 1993, 1996 2nd ed.

Koshland, D., "Fraud in Science" in: *Science* 235 (1987)

Lamboo, T., *Integriteitsbeleid van de Nederlandse Politie,* diss., Amsterdam: VU Amsterdam, 2005

Levelt, W.J.M. et al., *Interim Rapportage inzake door Prof. Dr. D.A. Stapel gemaakte inbreuk op wetenschappelijke Integriteit,* 31 October 2011, Tilburg: Tilburg University, 2011; second interim rapport, April 2012

Levelt, W.J.M. et al., *Falende wetenschap: De frauduleueze onderzoekspraktijken van sociaal-psycholoog Diederik Stapel.* Report by the Levelt Committee, issued 28 November 2012

Lewis, C.S., "The inner ring", in: *They Asked for a Paper: Papers and Addresses,* London: Geoffrey Bless 1962, pp. 146-147.

Lewis, C.S., *Studies in Words,* Cambridge: Cambridge University Press, 1960

Maas, P.J. van der, J. J. M. van Delden and L. Pijnenborg, *Medische Beslissingen rondom het Levenseinde,* The Hague: Sdu, 1991

Marks, J., "Scientific misconduct" in: J. Marks, *Why I am not a Scientist,* Berkeley: University of California Press, 2009, pp. 162–197

Medawar, P.B., *The Limits of Science,* New York: Harper and Row, 1984

National Academies of Science, *On Being a Scientist* (2012), 4e ed. NAS, Washington DC

NWO, *Reglement Wetenschappelijke Integriteit.* NWO, The Hague, 2005

Popper, K.R., *Objective Knowledge, an Evolutionary Approach,* Oxford: Clarendon Press, 1972

Protocol: Ethiek voor Wetenschappelijk Onderzoeken en Richtlijnen voor het Handelen bij (Vermeend) Wetenschappelijk Wangedrag, Leiden: Faculty of Social Sciences, Leiden University, 1998

Resnik, D.B., *The Ethics of Science, an Introduction,* London: Routledge, 1998

Rippen, K., "Naar eer en geweten, integriteit in het bedrijfsleven", in: F. Kok and T. van der Maas, *Zuiver op de Graat, hoe integer is Nederland,* Assen: Van Gorcum, 2006, pp. 61-90

Universiteit Leiden, *Protocol: Ethiek voor Wetenschappelijk Onderzoeken en Richtlijnen voor het Handelen bij (Vermeend) Wetenschappelijk Wangedrag*, Leiden: Faculteit der Sociale Wetenschappen RU Leiden, 1998

Voorbrood, C. and H. van Luijn, *Data – Voer voor psychologen: Archivering, beschikbaarstelling en hergebruik van onderzoeksdata in de psychologie*. Dans Studies in Digital Archiving 4. Amsterdam: Aksant Academic Publishers, 2010

VSNU*, Nederlandse Gedragscode Wetenschapsbeoefening*, The Hague: VSNU, 2004/2012

Weber, M., *Wissenschaft als Beruf,* Berlin: Duncker und Humblot, 1919 (Wetenschap als beroep. Politiek als beroep. Weber M. (transl.) H. Driessen, Nijmegen: Van Tilt 2012)

Williams, B., "Integrity" in: J.J.C. Smart and B. Williams, *Utilitarianism: For and Against,* Cambridge: Cambridge University Press, 1973, pp. 108-117

APPENDIX 1.
RESOLUTION INAUGURATING THE COMMITTEE ON SCIENTIFIC RESEARCH DATA

The Board of the Royal Netherlands Academy of Arts and Sciences (KNAW, "the Academy"), given Section 8 of the Regulations governing the Academy [*Reglement van de KNAW*] and considering that discussion has arisen within the world of research and within society in general regarding how researchers deal with research data, has resolved to appoint the Committee on Scientific Research Data, referred to hereinafter as "the Committee".

Section 1. Task of the Committee

The task of the Committee is to draw up recommendations encouraging researchers in all disciplines to familiarise themselves with routines that promote scientific integrity in their data management practices.

To that end, it will carry out a survey of the existing practices in various scientific fields for data collection and dissemination, including the standards applied by those directly involved.

The Advisory Report will also examine the institutional aspects of research data management. These involve, for example, the formal and informal responsibility of researchers and their employers for disseminating and complying with standards related to breaches of scientific integrity, for instance in teaching and supervising young researchers.

The Advisory Report will be drawn up in Dutch.[3] It is intended for researchers and administrators and Dutch public research institutions, including universities and the institutes that fall under the Netherlands Organisation for Scientific Research (NWO) and the Academy itself.

The Committee will ensure that the Advisory Report can be submitted to the Academy's Board before 1 April 2012.[4]

3 The Academy's Board ruled on 9 July 2012 that an English translation of the Advisory Report should also be published.
4 This date was later postponed to September 2012 at the request of the Academy's Board.

Section 2. Composition of Committee and Appointment Period

The following persons have been appointed (in their private capacity) to membership of the Committee:
Prof. C.J.M. Schuyt, chair
Prof. J.M. Bensing
Dr I.A.L. Stoop
Prof. J.P. Vandenbroucke
Prof. G.J. van der Zwaan
Prof. M.B.M. van der Klis
Dr P.K. Doorn will act as external adviser to the Committee.
The Committee will receive support from the Academy's Staff Department in accordance with the instructions of the Director General. The head of Policy Advice: Research and Knowledge Division (BWK) will act as the official secretary to the Committee.

The Committee will be appointed with effect from 1 November 2011 until 1 April 2012 (see footnote 3).

Section 3. Quality Assurance

Before being appointed, the members of the Committee will have familiarised themselves with the Preamble to the Advisory Committee Declaration of Interests [*Belangenverklaring Adviescommissie*] and completed and returned the Academy Advisory Committee Declaration of Interests Form [*Formulier Belangenverklaring Adviescommissie KNAW*].

Section 4. Costs and Allowances

Pursuant to Section 18(2) of the Regulations governing the Academy, the members of the Committee will receive a travel allowance.

Section 5. Confidentiality

The Committee will observe confidentiality in respect of all information that is made known in the context of the implementation of this resolution and that can be considered to be of a confidential nature.

Adopted in Amsterdam by the Board of the Royal Netherlands Academy of Arts and Sciences on 7 November 2011.

On behalf of the Board of the Royal Netherlands Academy of Arts and Sciences,

Dr K.H. Chang
Director General, Royal Netherlands Academy of Arts and Sciences

APPENDIX 2.
PEER REVIEWERS

The following persons carried out a peer review of this report at the request of the Board of the Royal Netherlands Academy of Arts and Sciences.

Prof. A.P. Buunk, University of Groningen
Prof. M.J.T.H. Goumans, Leiden University Medical Centre (LUMC)
Prof. E.P.J. van den Heuvel, University of Amsterdam
Prof. M.M. Levi, Amsterdam Academic Medical Centre (AMC)
Prof. F. van Oostrom, Utrecht University
Prof. J. Sixma, Utrecht University
Prof. G.H. de Vries, University of Amsterdam

APPENDIX 3.
INTERVIEWS AND WRITTEN REPLIES

The Committee conducted interviews with 15 persons on 20 January and 3 February 2012. These were eight professors (2 medical, 2 natural sciences, 4 behavioural and social sciences), 1 director of research (medical), 2 postdocs, members of The Young Academy (medical-biological, humanities), and 4 PhD candidates (2 medical, 1 technical 1 agricultural sciences). The names of the persons interviewed are:

Prof. F. Bovenkerk	Utrecht University, cultural anthropology
Prof. J.J.M. Braat	Delft University of Technology, optics
Prof. A. F. Cohen	Leiden University, Leiden University Medical Centre, Director, Centre for Human Drug Research,
Prof. C.K.W. de Dreu	University of Amsterdam, Psychology
Prof. N. Ellemers	Leiden University, Social Psychology
Prof. M.J.T.H. Goumans	Leiden University, Leiden University Medical Centre, cell biology
Prof. B.A. de Graaf	Leiden University, modern history
Prof. P.J.J. Hooykaas	Leiden University, biology
Prof. J.A. Knottnerus	VU University Amsterdam, General Medicine, chair of the Scientific Council for Government Policy (WRR)
Prof. P.J.M. Levelt	Max-Planck Institute for Psycholinguistics, cognitive development
A. Malan	Delft University of Technology, Leiden University Medical Centre, PhD candidate in information science in patient-related research, board member of Dutch PhD Candidate Network PNN (*Promovendi Netwerk Nederland*)
W. Ondracek	CEO Promasys BV (data management), Leiden
H. Pera	Wageningen UR, PhD candidate in physical chemistry, board member of PNN (*Promovendi Netwerk Nederland*)
B. Siegerink	Leiden University, Leiden University Medical Centre, PhD candidate in clinical epidemiology
J. Trietsch	Leiden University, Leiden University Medical Centre, PhD candidate in cancer research

The Committee drew up seven questions in advance that proved valuable in producing answers containing information. The Committee sent these questions to an estimated 130–150 persons,[1] 79 of whom replied. These included the chairs of the Academy's divisions, directors of research schools, and via them PhD candidates and postdocs. Members of the Committee also distributed the questions within their own organisations.

The distribution of replies according to scientific discipline and position was as follows: 36 professors, 16 directors of research/postdocs, and 27 PhD candidates representing the following fields:

Scientific field	**Number per field**	**Of whom professors**	**Of whom directors of research and postdocs**	**Of whom PhD candidates**
Earth sciences	7	1	1	5
Biology	8	4	2	2
Medicine	12	4	2	6
Mathematics, physics, and astronomy	7	5	0	2
Chemistry	5	4	1	0
Technical sciences	5	3	1	1
Agricultural sciences	5	2	1	2
Behavioural and social sciences	16	7	3	6
History	4	1	3	0
Philosophy, theology, law	5	3	2	0
Language and literature studies	5	2	0	3
Total	79	36	16	27

1 The precise number is unknown because PhD candidates were approached as a group at a number of research schools.