

**INTERIM REPORT REGARDING THE BREACH OF SCIENTIFIC
INTEGRITY COMMITTED BY PROF. D.A. STAPEL**

Tilburg, 31 October 2011

Contents

1	Introduction	3
1.1	Terms of reference	3
1.2	Coordinated inquiries at Tilburg, Groningen and Amsterdam.....	3
1.3	Working method of the Committee.....	4
1.4	Nature of the report	4
2	Inquiry findings	5
2.1	Extent of the fraud.....	5
2.2	Nature of the fraud	5
2.3	No signs of complicity of coauthors	6
2.4	Dissertations	6
2.5	Whistle blowers.....	7
2.6	Impact of the fraud	7
3	Research culture	8
3.1	Introduction	8
3.2	Mr Stapel’s methods.....	8
3.3	Failure of scientific criticism: peculiarities of Mr Stapel’s data and procedures.....	11
4	Recommendations	16
4.1	Further handling of the Stapel case	16
4.2	Code of conduct and supervision	17
4.3	Doctoral research.....	17
4.4	Psychology research practice and publications	18
	Appendix 1 Levelt Committee interviewees.....	19
	Appendix 2 Noort Committee interviewees	20
	Appendix 3 Mr Stapel’s response to the draft interim report	21

1 Introduction

1.1 Terms of reference

The Rector Magnificus of Tilburg University, Prof. P. Eijlander, formed the Levelt Committee on Friday, September 9, 2011. The members of the Committee are Prof. W.J.M. Levelt (chair), director emeritus of the Max Planck Institute for Psycholinguistics and former president of the Royal Netherlands Academy of Arts and Sciences, Prof. M.S. Groenhuijsen, full professor of criminal law, criminal procedure and victimology and founder-director of the interfaculty research institute Intervict, and Prof. J.A.P. Hagenaars, emeritus professor of Methods and Techniques of Social Scientific Research and former dean of the Faculty of Social Sciences. The Committee secretary is Dr S.A.M. Baert.

The Rector Magnificus gave the Committee the task of investigating the extent and nature of the breach of scientific integrity committed by Mr D.A. Stapel. There are two elements to the task:

1. the committee will examine which publications are based on fictitious data or fictitious scientific studies and during which period the misconduct took place;
2. the Committee should form a view on the methods and the research culture that facilitated this breach, and make recommendations on how to prevent any recurrence of this.

1.2 Coordinated inquiries at Tilburg, Groningen and Amsterdam

The inquiry extends beyond Tilburg University, in view of Mr Stapel's former appointments at the University of Amsterdam and the University of Groningen. Both these other universities have formed committees of inquiry. The members of the Amsterdam committee are Prof. P.J. Drenth (chair, VU University Amsterdam), Prof. J.W. Zwemmer, Prof. L.A. de Klerk and Prof. C.A.J. Klaassen (all of Amsterdam University). The members of the Groningen committee are Prof. E. Noort (chair), Prof. H.E. Bröring (both of the University of Groningen), and Prof. J.M. Pieters (University of Twente). The administrative secretary is *drs* M. Jaspers. The three committees cooperate and consult each other, and the Levelt Committee of Tilburg University coordinates their work. The distribution of tasks is as follows: the University of Amsterdam committee investigates the publications from 1994 to 1999, inclusive, with particular attention to the research on which Mr Stapel's dissertation was based; the Noort Committee of the University of Groningen investigates the publications from 2000 to 2006, inclusive; and the Levelt Committee of Tilburg University investigates the publications from 2007 to 2011, inclusive. Regarding Mr Stapel's PhD students and those whose PhD was awarded under his supervision, the Groningen committee investigates the dissertations defended in Groningen, and the Tilburg committee those in Tilburg.

The three committees report to their respective Rectors or Executive Boards. As agreed, any specific Amsterdam and Groningen contributions, deviations, additions, and corresponding findings have been inserted into this interim report and marked (AMS and GRO, respectively) at the appropriate places in the text.

1.3 Working method of the Committee

The Committee started by collecting all Mr Stapel's publications published in 2007 to 2011, inclusive, including all datasets, questionnaires, hypotheses and e-mail correspondence.

Mr Stapel himself was also requested to provide a list of publications that were produced using fictitious data. He provided a list of his journal articles from 1994 up to and including 2011, with an indication of any that were based on fabricated data. Regarding a second list, referring predominantly to chapters in books and proceedings, Mr Stapel has stated that he is physically and emotionally unable to respond. Furthermore, many interviews were held with whistle blowers, present and former PhD students, coauthors, colleagues and present and former members of the faculty board, heads of the department and research directors in order to build up the clearest possible picture of the course of events. A full list of those consulted is given in Appendix 1. The list of those consulted by the Groningen committee is given in Appendix 2. The Amsterdam committee's consultations were still ongoing at the time this interim report was released.

The Committee has upheld the principle of hearing both sides of the case. Mr Stapel was invited for a personal interview with the Committee. Unfortunately, he indicated that he was unable to comply. A draft of this interim report was accordingly sent to him and his lawyer for comment. Mr Stapel's comment on the draft interim report is attached as Appendix 3.

GRO: The Noort Committee in Groningen took the same line for the 2000-2006 period. Also considered were the students who obtained their PhD in Groningen after this time and where Mr Stapel had a role as supervisor. In the first instance there was no differentiation according to the order of rank of the coauthors. A questionnaire focusing on the manner of conducting experiments, and the acquisition, entry, analysis and interpretation of the data, was sent to all (35) coauthors in the relevant period. Some were then invited for interview, depending on the answers and their position in the order of rank of authors. Where possible these interviews were conducted face-to-face, and otherwise by Skype. In Groningen too, colleagues and present and former members of the faculty board, heads of the department and research directors were invited in order to build up a picture of the research climate.

1.4 Nature of the report

The purpose of this interim report is to give an indication of the extent of the fraud, to describe the nature of the fraud, and the research culture within which it took place. The Committee has come to the conclusion that further investigation will be needed in order to produce a full list of publications in which use was made of fictitious data. However, the many interviews have enabled the Committee to obtain a satisfactory picture of the nature of the fraud and the research culture in which the use of fabricated data has persisted for so long. It has now been possible to formulate a number of concrete recommendations based on this picture.

However, it cannot be ruled out that additional recommendations will appear in the final report, and it may be possible to add details of the fraud and what allowed it to occur. These further recommendations and details will come mainly in the final reports

of the Amsterdam and Groningen committees, which started their work somewhat later than the Tilburg committee. The final report, too, will be drawn up as a coordinated effort.

2 Inquiry findings

2.1 Extent of the fraud

The Committee has arrived at the conclusion that the extent of Mr Stapel's fraud is very substantial. The Committee has already encountered several dozen publications in which use was made of fictitious data. The full list of these fraudulent publications will be presented jointly with the other committees in the final report, or before its release. It has been established that the fabrication of data started even before the Tilburg period. The Committee has concluded that publications from the Groningen period are also affected. The fraud has persisted for some considerable time: at any rate since 2004.

GRO: In his Groningen period (2000-2006), Mr Stapel also engaged in fraudulent data manipulation. As far as the committee has been able to establish, several of these data manipulations were performed with data ostensibly collected in his Amsterdam period. Also in his Groningen period, datasets were regularly fabricated, the analysis manipulated, or both.

2.2 Nature of the fraud

The Committee has established that Mr Stapel has committed fraud with data. He has at any rate fabricated datasets in several dozen studies. The preparation for these studies proceeded normally. Stimuli and questionnaires were developed following comprehensive discussion. In the end, however, the questionnaires were not administered. Mr Stapel created his own datasets, which generally confirmed the expectations. In some cases, Mr Stapel would send fellow researchers datasets that he claimed to have had in his possession for some considerable time, asking them to analyze and write a first draft of the paper. Datasets appear to have been fabricated in these cases as well. A comprehensive description of Mr Stapel's methods is given in Section 3.2.

The fact is that the fraud with data has been on a large scale and has persisted for a lengthy period, so that people, and in particular young researchers entrusted to him, have been affected profoundly at the start of their careers. This conduct is deplorable, and has done great harm to science, and the field of social psychology in particular. To the best of our knowledge, misconduct of this kind by a full professor in his position is unprecedented.

GRO: The Noort Committee has established that Mr Stapel has committed fraudulent data manipulation in various ways. The fraud involved creating data ostensibly from experimental subjects, who ultimately appeared not to exist. For example, Mr Stapel would claim that a questionnaire had been distributed in a school or among a group of experimental subjects, but the school or subjects did not exist, or proved impossible to trace after the event. There is evidence of three forms of fraud: (1) the complete invention of data; (2) the fabrication or augmentation of a database after Mr Stapel's initial inspection gave him the impression that the data would not yield what he

expected (confirmation of hypotheses); and (3) modification of complete or incomplete data. There is a strong suspicion that these three forms occurred in the methods Mr Stapel used to create pliable datasets. This suspicion is supported by Mr Stapel's regular gesture of offering PhD students or colleagues a dataset, which he claimed to have had in his possession since his Amsterdam period, but had not found time to analyze. Further investigation by statisticians will help in assessing these suspicions.

2.3 No signs of complicity of coauthors

The Committee has as yet found no signs that coauthors knowingly cooperated with Mr Stapel in tampering with data, nor that they were aware, or could or should have been aware, of any tampering with data.

2.4 Dissertations

The Committee has established with certainty that fictitious data were also used in doctoral research supervised by Mr Stapel in Tilburg. The Committee first established that the doctoral dissertation of Dr F. van Horen is not based on fictitious data. On the other hand, it is certain that one or more chapters in the dissertations of Dr Y.R. Avramova, Dr M.K. Noordewier and Drs M.H.C. Meijers do rely on fictitious data. The same may apply to subsidiary parts of the dissertations of Dr L. van den Broek and Dr C. Wiekens. Ms Meijers was due to defend her doctoral dissertation on October 7, but the ceremony has been postponed at her request upon learning of the data fraud, because the dissertation may also contain 'tainted' data.

The Committee then asked whether these findings should have repercussions for previously awarded degrees. The Committee used the following criterion: were the PhD students concerned aware that the data were fictitious, or was there an element of culpable ignorance? In answer, the Committee has ascertained that none of the cases involved culpability, nor complicity on the part of the PhD students in committing fraud. On the contrary, all these PhD students and coauthors were misled with great subtlety. The Committee is therefore of the opinion that there should be no repercussions on the awarded degree in any of these cases. Nonetheless, it is proper that the bias in the data used in these dissertations be made known, in order to prevent possibly false conclusions being drawn, and to curtail the impact of interpretations that have been made.

GRO: Following discussion with all researchers whose PhDs were earned in Groningen, the Noort Committee arrived at the same conclusion as the Levelt Committee about articles and chapters published prior to, or based on, their dissertations. There was no question of any knowledge or cooperation in the fraud on the part of these PhDs. Likewise, the concept of culpability as introduced by the Levelt Committee does not hold for these junior researchers in Groningen. The Noort Committee therefore endorses the Levelt Committee's opinion that the data fraud committed by Mr Stapel should have no repercussions on awarded degrees.

The Noort Committee used a more detailed operational definition regarding the dissertations, as distinct from the people who were awarded the PhDs. The criterion used by the Noort Committee was therefore to what extent data was collected, input, analyzed and interpreted by the researchers themselves. In some cases datasets provided by Mr Stapel had been used. There were also cases in which data at a certain

stage of analysis and interpretation were in Mr Stapel's possession outside the control of the PhD student.

The Noort Committee has no suspicions regarding the doctoral dissertations of Dr M.B. Braun-Ekker, Dr E. Kamans, Dr S.J. Ko, Dr J. Lammers, Dr H. Oldenhuis, and Dr S. van der Velde. In these cases Mr Stapel was a supervisor for formal reasons only, or it has been established that the PhD student had control of all data from start to finish. The situation is different for the dissertations of Dr Janneke F. Jolij, Dr J. Grob, Dr S.A.Schwinghammer, Dr D. Trampe, Dr A. van den Bos, Dr M. Maringer and Dr L. Renkema. In these cases the Noort Committee has reservations, especially regarding the acquisition and personal management of parts of the data.

2.5 Whistle blowers

Three young researchers from the department concerned voiced their suspicions about the forgery of data by Mr Stapel to the head of department at the end of August 2011. After months of observation, sufficient details had been gathered to demonstrate that something was not right. The researchers all deserve praise for reporting these abuses. It is noted that they were in a dependent position and had much to lose. Following the report, the head of department immediately informed the Rector Magnificus, who in turn spoke immediately with the whistle blowers. It is beyond the Committee's remit to describe, or to evaluate in detail, the evidently effective subsequent action of the Rector Magnificus and university governing board.

The Committee has learned that there were other whistle blowers at an earlier stage. A total of three young researchers had previously raised the alarm regarding datasets provided by Mr Stapel, and the risks for them were no less severe. Furthermore, two professors had previously observed data that were 'too good to be true'. However, none of these earlier reports were acted upon. There were various reasons for inaction, which we will return to in Section 3.3.

2.6 Impact of the fraud

The Committee finds that great harm has been done to coauthors, and PhD students in particular, as a consequence of the fraud committed by Mr Stapel. The Committee has become aware in the course of the interviews that the consequences for those involved can be both formal and informal in nature. In a formal sense, the people affected are hampered in their careers, such as when extending temporary contracts and applying for grants. On an individual level some victims have been demoralized. In an informal sense, there is an element of stigmatization that may persist long into their further career.

It is not only individual people that are affected, but confidence in science, and in social psychology in particular, has been badly dented by Mr Stapel's actions. Other victims are the universities of Tilburg, Groningen and Amsterdam, whose reputations have been damaged, academic publishers, which have been obliged to withdraw published articles, providers of research funds, whose grants have been used fraudulently, and fellow researchers, who may have been denied the grants and direct funding that were awarded instead to Mr Stapel.

3 Research culture

3.1 Introduction

What were the factors that allowed the data fraud to acquire this shocking scale and persist for so many years? An answer to this question is needed to support recommendations for preventing recurrence that are to be presented in Part 4.

As concluded above in Part 2, there are no signs that anyone else knowingly cooperated in the data fraud. On the contrary, the Committee is convinced after all the interviews and inquiries that the PhD students and other coauthors must be considered victims of Mr Stapel's fraudulent methods. How could Mr Stapel persist in these methods for so long? This question is all the more acute in view of the conspicuous nature of various aspects of Mr Stapel's methods, even with no direct suspicion of fraud, which had indeed been remarked upon as unusual by some people.

The answer has two parts. The most important reason is the sophistication of Mr Stapel's methods, and the unscrupulous use of his prestige, reputation and influence. This aspect is explained in Section 3.2. A second relevant factor was the flawed performance of academic criticism, which is the cornerstone of science. It was easy for Mr Stapel's methods to take advantage of this situation. This aspect is the subject of Section 3.3.

GRO: The Noort Committee's findings on the research culture largely concur with those of the Levelt Committee. The points of agreement include Mr Stapel's charisma and unassailable position within the research group, and the great trust placed in Mr Stapel's scientific knowledge, expertise, and integrity by PhD students and close colleagues. However, the findings of the Noort Committee show a different emphasis. In the light of the investigation so far, the emphasis in the Groningen period of Mr Stapel's fraud appears to have been less on inventing research data than on massaging them (see also Section 2.2).

3.2 Mr Stapel's methods

Reciprocal trust is an essential part of all teamwork, also in scientific research. This is pre-eminently the case with the kind of master-apprentice relationship that occurs in science (e.g. supervisor-PhD student, or senior researcher-junior researcher). The junior must be able to trust absolutely in the integrity of the 'master'. For everyone who collaborated academically with Mr Stapel, including in particular his peers and juniors in the Department of Social Psychology, trust in the scientific integrity of Mr Stapel was absolute. The last thing that colleagues, staff, and students would suspect is that, of all people, the Department's scientific star, and Faculty dean, would systematically betray that trust. Many saw him as a charismatic leader whose dedication to staff and students was great. Only few would have had the courage or inclination to consider data fraud.

Mr Stapel employed a cunning, simple system in his data fabrication. The essence of the commonest form was as follows. He would develop, in intensive one-to-one interaction with another scientist (usually a research Master's student, a PhD student, or post-doctoral researcher, but sometimes a senior colleague), a theory that conformed intelligently and creatively with a research question, which usually came from the

partner. Ingenious experimental manipulations were designed to test the theory, which were prepared and documented in detail. Mr Stapel would demonstrate substantial knowledge of the literature, and considerable skill in designing appropriate experimental manipulations. After careful discussion, the materials for the experiment, such as precise descriptions of the experimental set-up, detailed instructions for the investigators on site, the piles of questionnaires, experimental supplies, such as pictures, sweets as reward, etc. would be prepared by the scientific partner in the quantities needed. Often, at Mr Stapel's request, the scientific partner would also produce a table of expected empirical findings of all the hypotheses to be tested. The materials would all be delivered to Mr Stapel, in many cases being placed in the trunk of his car, for transport to the school concerned.

Up to this point everything was in order. However, there then followed an entirely fictitious phase. The experiments were seemingly executed under the complete supervision of Mr Stapel alone. According to his own account, Mr Stapel had excellent contacts with many educational institutions in the country. They were apparently always willing to be persuaded by him to perform studies of this kind, sometimes with the assistance of (fictitious) paid research assistants. To compensate the schools for their efforts, Mr Stapel claimed to give lectures and provide the schools concerned with occasional computer time and the use of video projectors. The data collected at these schools were then said (falsely) to have been processed and coded at the schools themselves, usually by (unknown) assistants. The data 'obtained' in this way were then given directly to Mr Stapel: never to the partners.

All these 'efforts' usually resulted several weeks later in a (fictitious) data file being made available in its entirety to the partner for more detailed analysis, or even directly in the form of tables with the necessary averages, standard deviations, reliability checks, test results, etc. The partner would be able to start writing the article immediately, which might involve intensive cooperation with Mr Stapel.

This method applies to most of the fraud encountered by the Committee. There were also variants of the above. Sometimes Mr Stapel would discuss with a colleague an interesting study that the colleague was working on, remarking that he had an old data set, which he had once collected, but had then done nothing with, and which matched the colleague's needs perfectly. Soon this fictitious data set would be made available to the colleague, who would then be in a position to produce a fine article, obviously with Mr Stapel as coauthor. He would also say in Groningen or elsewhere that he had obtained data from the Social Psychology lab in Tilburg, or could arrange for the data to come from Tilburg. The data involved could then be fictitious, as they were in the disputed 'meat research'. In Tilburg, the same could happen in reverse with respect to Groningen. Also, at least once, the original dataset was withdrawn by Mr Stapel for more detailed analysis in response to comments from a journal. The data appeared later to have been augmented with some ten new observations, as a result of which all the hypotheses were neatly confirmed. Finally, some entire studies were inventions, and 'executed' in solo, e.g. a study in a train station with unknown research assistants.

The fraud was dressed up with great subtlety. The most important obfuscation was the many weeks, or even months, of painstaking preparation, leaving no room for doubt that the research was real. Realistic details were also added, such as that account had been taken in the conduct of the experiments with exam periods and school vacations. Despite the vague information about the schools, verifiable details were provided about

the nature, location and specific curricular and other characteristics (which were actually verified on Internet). It was said that research in schools was important because school students, unlike the usual psychology students, were more naive and therefore better research subjects. The schools were to be approached by Mr Stapel alone, not by other researchers, to prevent the schools being overrun with similar requests, which would hamper Mr Stapel's access to them. Mr Stapel justified his exclusive access to data on the lab computer by invoking the risk of introducing a virus if they were available to everyone (although other professors tended not to adopt a similar approach). More could be added to this list, but it gives a sufficient impression of the 'ingenuity' with which the fraud was planned and set up.

AMS: The data from the Amsterdam period was based less on studies among high school students than in later research, although this kind of data was used, among other things, for some articles on which his doctoral dissertation was based. In most cases, the data came from studies involving first or junior year students (Amsterdam, Eindhoven), or undergraduate students (University of Michigan, University of Georgia). First-year psychology students at the University of Amsterdam thus fulfill an obligation to act as experimental subjects for a number of hours. Students sometimes also took part in studies voluntarily or for payment. In the years concerned research was indeed conducted on first-year students, and Mr Stapel participated. Here too, Mr Stapel nearly always took charge of collecting, coding and processing the data in isolation and without being checked or verified. The data are no longer available, having been destroyed in accordance with the Netherlands Code of Conduct For Scientific Practice. It is, therefore, impossible to check whether these data are affected by fraud. The same (procedure and availability of the data) applies to the 'American' data.

Mr Stapel carefully selected the junior researchers with whom he embarked on the intensive one-to-one method. A close bond was formed with them. Many PhD students viewed him as a personal friend. They would visit him at home, and they would dine together, go to the cinema together, etc. This kind of situation is not conducive to following the 'master' critically.

With the method outlined above, Mr Stapel bears full responsibility for the enormous harm inflicted on people and their careers, and on the reputation of science, and social psychology in particular.

GRO: The impression gained by the Noort Committee is that when a reviewer or editor made a request as part of a review process for supplementary analysis or additional data, the items requested were delivered in the desired form and outcome relatively rapidly and by vague means. These forms of fraud were therefore possible because Mr Stapel's methods (see also Section 3.2) were subtle and exceptional, and might have raised suspicion had it not been for his convincing yet intimidating attitude, which permitted no other implementation. Mr Stapel took charge of processing the data. This was plausible because of the many grants he had attracted, thereby allowing him to afford a series of research assistants. Mr Stapel's researchers, PhD students and colleagues tended to gain access to the data only after Mr Stapel had 'inspected' them following processing. In many cases, the SPSS output was given to the researcher, who was able rapidly to incorporate the findings into an article, with no need to spend much time on processing, and in some cases analyzing, the data.

GRO: Despite the occasional difference in emphasis in Groningen, the essence of Mr Stapel's actions was almost identical. Within his research group Mr Stapel was the absolute 'master of the data'. Whereas in other groups the conduct of the experiments, the entry of the raw data, the analysis and interpretation were also used as learning opportunities, where Mr Stapel was concerned, one or more phases sometimes ran via him only. An assistant entered the data from a study into the computer and then gave the file to Mr Stapel. A copy of the file was rarely sent directly to the researcher. The researcher would receive the file for analysis from Mr Stapel, with all the attendant opportunities for interim manipulation. More than once Mr Stapel would have performed some preliminary work. 'Be aware that you have gold in your hands' the young researcher would be told. An acute difference of opinion arose with one young researcher, who persistently pressed for access to the raw data. Mr Stapel suggested that the researcher was calling his capacities and experience as a renowned professor into question. PhD students and junior researchers did comment on Mr Stapel's control over the data, and also discussed it with colleagues, but ultimately viewed it as a service. Mr Stapel enjoyed large research budgets, was able to appoint many student and other assistants, and could thus save the young researchers time that would otherwise have been spent on analysis, interpretation and writing. Moreover, the procedure sometimes provoked the envy of young researchers who could not benefit from this service.

3.3 Failure of scientific criticism: peculiarities of Mr Stapel's data and procedures

As cleverly as the fraud was set up, various characteristics of the methods used by Mr Stapel and the data he generated could have aroused incredulity, which they indeed did with various people. Why then did it take so long for the fraud to be discovered?

The reason is not only that this unprecedented fraud was unthinkable, but also has to do with a failure of rational, systematic, substantial, methodological and public criticism, which is the cornerstone of science.

Given below are various peculiarities in the data collection and the actual data that should have been observed in a normal vital scientific world, and indeed were sometimes remarked upon, albeit without further action. The parties who could have been in a position to do something about the matter were research Master's students, PhD students, doctoral boards, post-doctoral researchers, junior and senior researchers and colleagues, staff and governing board of the Department, the Faculty, or the University, and the editorial teams and reviewers of journals. The analysis of this failure of criticism in no way implies that the people or bodies involved were complicit in the fraud. At the present state of the inquiry, the fraud is solely and entirely the work of Mr Stapel. However, his practice of data fabrication would certainly have been more difficult, if not impossible, had there been other constraints on research conduct on multiple levels. The purpose of our analysis is to draw lessons from this lamentable actual case, and Part 4 of this interim report makes recommendations that may lead to better protection of scientific integrity, in particular in social psychology and psychology research.

First, then, a number of peculiarities surrounding the fraudulent data collection are given below.

Mr Stapel (claimed to have) collected the data himself, had sole supervision of the experiments, and processed and coded all the data either himself, or using unknown assistants under his direct supervision.

It is important for a PhD student or research Master's student to gain personal experience of the entire research process, including the collection and processing of the data, and certainly so where their own research is involved. A number of Mr Stapel's PhD students therefore never experienced this process for themselves. It is strange in this respect that, although this principle is unanimously endorsed by his colleagues in the Department, no one knew that Mr Stapel, and not the PhD students, collected the data. Assuming a joint responsibility of the Department for education, also of research Master's students and PhD students, this is surprising. Apparently, Mr Stapel was able in practice to do whatever he wanted, with no systematic, structural consultation.

Had the general rule in the Department been to have two supervisors (one day-to-day coach and a supervisor) for each PhD student, this form of fraud would have been more difficult. Fraud would also have been more difficult if PhD students had worked in teams. As it was, researchers selected by Mr Stapel operated exclusively in this bilateral, isolated relationship. With more teamwork and consultation between PhD students, the fact that a PhD student collected no data would have been much more conspicuous.

In that respect, the responsibility of the Graduate School, under whose faculty responsibility the research Master's and PhD education resides, did not fully materialize.

Various PhD students, after their intensive preparation, asked to be allowed to lead the experiment themselves, or to see it in action. The requests were always dismissed with an excuse: they could make better use of their time for the real scientific work (analyzing and writing); the schools were also adamant that no-one except Mr Stapel should contact or visit them. Requests were sometimes also made for access to the completed questionnaires. Unfortunately this proved impossible, since the questionnaires had already been coded and entered into the computer at the schools. Neither the schools nor Mr Stapel himself had enough space to store all those piles of paper. Any probing questions were usually cut short with an appeal to the trust that Mr Stapel was entitled to.

It was precisely because of the isolated approach that the young researchers were unaware that this was not a normal state of affairs in social psychology research.

Regarding the collaboration with senior researchers, Mr Stapel also took personal charge of the 'data collection' and provided the outcomes, but not the raw data. This distribution of tasks, based on trust, is not unusual in science.

The members of the doctoral boards where fraudulent data were involved were also never aware that the PhD students usually did not collect the data themselves, but were supplied with them by Mr Stapel. PhD students wrote with full conviction (in line with recommendations given in the Manual of the American Psychological Association (APA)): we collected the data at schools etc., using the collective 'we' in a way viewed as entirely legitimate and usual by the PhD students and the members of the doctoral boards alike. However, if the rule had been for the doctoral boards to ascertain that the

data had been collected in principle by the PhD students themselves, then this data fraud could not have worked.

Insufficient disclosure and specification of the method; the experiments were too complex for the schools; there were no completed questionnaires; schools were never identified.

Many aspects of the data collection process were unclear. Journals sufficed with references in the articles such as ‘Universities in the North of the Netherlands’. The precise identification of populations and samples is nonetheless essential. As stated above, PhD students and other partners who asked for the names of the schools were given vague descriptions, with which they usually sufficed. Anyone asking probing questions would be given more details, but almost never the exact locations of the schools, and always accompanied with a prohibition to make contact with these schools. Those concerned were expected to trust Mr Stapel. Anyone who became too critical ran the risk of the intensive bilateral collaboration being terminated, or the suitability for a possible PhD student position being called into question.

GRO: The Noort Committee observed that (1) experimental subjects were described as ‘undergraduate students’ or ‘participants’, with no further statement of origin. In a few cases it was noted that the ‘undergraduates’ were from Groningen. (2) When PhD students were unaware of the origin, ‘undergraduates’ was stated, although they were probably pre-university school students.

Senior partners generally did not get involved in these kinds of (methodological) matters, but should have been concerned about the lack of explicitness.

The critical stance also failed in respect of the actual feasibility of the experiments in the schools. Some requirements demanded comparisons between buckets of water held at constant temperature. How can that be achieved in a classroom? The application of various projective techniques and observations requires intensive training of the investigators (teachers?) and sometimes one or two investigators for each child. Is it reasonable to expect schools to be able to do this repeatedly? However, no questions were asked, and it was simply accepted that Mr Stapel somehow managed to make it happen.

Apparently neither the reviewers nor the editorial teams of journals delved into aspects of this kind. The lack of space in the articles themselves understandably prevents comprehensive coverage of these aspects of research, but questions should still have been raised about the actual course of events in the reported study. It is no unnecessary luxury for journals to demand a detailed report of the research procedures followed to be made available on Internet.

The data were too good to be true; the hypotheses were almost always confirmed; the effects were improbably large; missing data, or impossible, out-of-range data, are rare or absent.

This is possibly the most precarious point of the entire data fraud. Scientific criticism and approach failed on all fronts in this respect. The falsification of hypotheses is a fundamental principle of science, but was hardly a part of the research culture surrounding Mr Stapel. The only thing that counted was verification. However, anyone

with any research experience, certainly in this sector, will be aware that most hypotheses that people entertain do not survive. And if they do, the effect often vanishes with replication. The fact that Mr Stapel's hypotheses were always confirmed should have caused concern, certainly when in most cases the very large 'effect sizes' found were clearly out of line with the literature. Rather than concluding that this was all improbable, instead Mr Stapel's experimental skills were taken to be phenomenal. 'Too good to be true' was meant as a genuine compliment to his skill and creativity. Whereas all these excessively neat findings should have provoked thought, they were embraced. If other researchers had failed, that was assumed to be because of a lack of preparation, insight, or experimental skill. Mr Stapel became the model: the standard. Evidently only Mr Stapel was in a position to achieve the precise manipulations needed to make the subtle effects visible. People accepted, if they even attempted to replicate the results for themselves, that they had failed because they lacked Mr Stapel's skill. However, there was usually no attempt to replicate, and certainly not independently. The few occasions when this did happen, and failed, were never revealed, because the findings were not publishable.

In other words, scientific criticism has not performed satisfactorily on this point. Replication and the falsification of hypotheses are cornerstones of science. Mr Stapel's verification factory should have aroused great mistrust among colleagues, peers and journals.

As a supervisor and dissertation advisor, Mr Stapel should have been expected to promote this critical attitude among his students. Instead, the opposite happened. A student who performed his own replications with no result was abandoned to his fate rather than praised and helped.

Strange, improbable, or impossible data patterns; strange correlations; identical averages and standard deviations; strange univariate distributions of variables.

The actual data displayed several strange patterns that should have been picked up. The patterns are related to the poor statistical foundation of Mr Stapel's data fabrication approach (he also tended to make denigrating remarks about statistical methods). It has emerged that some of the fabrication involved simply 'blindly' entering numbers based on the desired bivariate relationships, and by cutting and pasting data columns. This approach sometimes gave rise to strange data patterns. Reordering the data matrix by size of a given variable sometimes produces a matrix in which one column is identical to another, which is therefore the simple result of cutting and pasting certain scores. It was also possible for a variable that would normally score only a couple of per cent 'antisocial', for no reason and unexpectedly suddenly to show 'antisocial' most of the time. Independent replication yielded exactly the same averages and standard deviations. Two independent variables that always correlated positively, conceptually and in other research, now each had the 'right' expected effects on the dependent variable, but correlated negatively with each other. There was no consistent checking of data by means of simple correlation matrices and univariate distributions. It is to the credit of the whistle blowers that they did discover the improbabilities mentioned above (see below).

Finally, a lamentable element of the culture in social psychology and psychology research is for everyone to keep their own data and not make them available to a public archive. This is a problem on a much larger scale, as has recently become apparent.

Even where a journal demands data accessibility, authors usually do not comply (Wicherts et al. 2006). Archiving and public access to research data not only makes this kind of data fabrication more visible, it is also a condition for worthwhile replication and meta-analysis.

Many of the peculiarities mentioned were revealed clearly only with the knowledge now available. However, people have repeatedly had a direct view of one or other aspect. Some discounted any misgiving as a stray thought, because they did not believe it, or felt powerless against Mr Stapel. However, there were others who gave the matter more attention, and yet others discussed it. Why did this discussion not lead to earlier discovery?

One possible factor was the late recognition of suspicions of fraud. In the past year, six young researchers have expressed doubts about Mr Stapel's scientific integrity. Who should you approach with doubts of this kind? In accordance with the rules of the National Committee For Scientific Integrity (LOWI), the person to approach is a Confidential Counselor appointed by the university. Contrary to LOWI guidelines ('The function of Confidential Counselor is incompatible with that of member of the Executive Board, faculty dean, director of a research school, or educational or research institute'), the Confidential Counselor in Tilburg was the Rector Magnificus. Certainly for a student or young PhD student, that is a very lofty and distant position. The Rector Magnificus is one of the most powerful people in the university. He is not someone you approach, it would be obvious to think, without watertight proof of your suspicions. Furthermore, since his appointment, and as dean, Mr Stapel had enjoyed considerable support from the Executive Board.

You would therefore approach a professor, or another member of staff, in the hope that they will know what to do next. And even this course of action is not without problems, since the kind of relationship they themselves have with the alleged fraudster is unknown. It makes it much easier for whistle blowers if there is an independent university confidential counselor for scientific integrity, who is not part of the management line.

GRO: The conspicuously unusual treatment of data was already a subject of discussion in the Groningen period, but, surprisingly, not among the direct colleagues of Mr Stapel's chair group.
--

On three occasions in 2010 and 2011, the attention of members of the academic staff in psychology was drawn to this matter. The first two signals were not followed up in the first or second instance. Mr Stapel's virtually unassailable position may have played a role. On the third occasion, a report was made to the head of department, which was extraordinarily carefully prepared by three young and, certainly in their position, vulnerable whistle blowers. This report was immediately picked up in a professional way, with the now familiar result.

Suspicions about data provided by Mr Stapel had also arisen among fellow full professors on two occasions in the past year. These suspicions were not followed up. The Committee concludes that the six young whistle blowers showed more courage, vigilance and inquisitiveness than incumbent full professors. The Committee nonetheless adds that at no time was any official within the university aware of compelling signs of academic misconduct that would have obliged him or her to devote

official attention to the matter.

4 Recommendations

Trust remains the basis of all scientific collaboration. It cannot be replaced by bureaucratic measures. However, this trust can thrive only in a research environment where the rules of the game are clear and explicit. On that point, in particular in psychology research practice and the handling of the Stapel case, the following recommendations must be observed.

4.1 Further handling of the Stapel case

It stands to reason that the universities of Groningen and Tilburg should be generous vis-à-vis the individuals whose honor and careers have been harmed.

Past and present PhD students

In the case of Drs Meijers, who has yet to be awarded her doctorate, the Committee recommends Tilburg University to give her an opportunity to perform supplementary research and to produce a revised doctoral dissertation based on untainted research material.

The university concerned should publicly disclose all cases in which a dissertation appears to have been based fully or partly on fabricated data.

GRO: It is worthy of consideration to issue PhDs with an official statement to the effect that, following thorough investigation, the Committee has absolved them of all blame, and that their degree was rightly awarded in the light of their own performance.
--

Coauthors

Coauthors of tainted articles must be informed as soon as possible of the status of the article concerned. Immediately upon completion of the work of the three cooperating committees of Tilburg University, the University of Groningen and the University of Amsterdam, the three universities should issue a joint list of 'tainted' Stapel (et al.) publications to the publishers of the articles concerned. They should also inform publishers of articles where no certainty can be obtained about the veracity of the data they report. Both lists must also remain publicly available, and likewise any subsequent updates.

Research funding bodies

The universities concerned must inform the providers of research funds for Mr Stapel's research about any fraudulent publications that have flowed from this research. They must also determine how the grants were actually used on Faculty and Department levels.

University funds

The universities concerned must determine whether proper use of university funds was made within Mr Stapel's research group.

Criminal report

The Committee recommends Tilburg University to inform the Public Prosecution Service of forgery of documents and fraud committed by Mr Stapel. An important aspect in this connection is the serious harm inflicted on the reputation and career opportunities of young scientists entrusted to Mr Stapel.

Investigation into stripping Mr Stapel of his doctorate

Regarding Mr Stapel's doctoral dissertation, the Amsterdam committee has been unable demonstrably to establish any fraud with research data, partly because the data has been destroyed. It is therefore legally impossible to revoke the doctorate on these grounds. However, the Committee recommends the University of Amsterdam, pursuant to the findings, to investigate whether Mr Stapel can be stripped of his doctorate on the grounds of exceptional academically unworthy conduct, in breach of the obligations attached to the doctorate. (<http://www.lto.de/de/html/nachrichten/4297/vgh-bw-entzieht-dokortitel-der-unwuerdige-forschungsbetruerger>).

4.2 Code of conduct and supervision

The Committee recommends that newly appointed members of Tilburg University staff be required to sign a contract obliging compliance with code of conduct. On commencement of their work PhD students must take a short integrity course covering the code of conduct. Recommended codes of conduct are 'The Netherlands Code of Conduct For Scientific Practice' (Association of Universities in the Netherlands, 2005) and the 'European Code Of Conduct For Research Integrity' (ESF-ALLEA 2011).

Tilburg University must appoint a Confidential Counselor For Academic Integrity in accordance with national guidelines agreed in LOWI. Every student and member of staff must be informed of the existence of this Confidential Counselor.

In addition to the current 'Tilburg University Arrangements For Whistle Blowers' (2007) Tilburg University – in accordance with national guidelines agreed in LOWI – must draw up and introduce rules for protecting whistle blowers when reporting a probable infringement of scientific integrity.

4.3 Doctoral research

Every doctoral research project in psychology at Tilburg University must have at least two supervisors.

The data on which an empirical psychology doctoral dissertation in Tilburg are based must as a rule be collected and analyzed by the PhD student concerned. The doctoral board is responsible for ascertaining compliance.

The Tilburg University doctoral formula must state both the rights and obligations towards science and society that are attached to the doctorate by law or custom.

AMS: The Doctoral Committee of the University of Amsterdam is recommended to investigate whether the above recommendations lead to any revision of the General Doctoral Regulations, in particular the articles dealing with the responsibilities of the
--

PhD student, supervisor and cosupervisor.

4.4 Psychology research practice and publications

The Committee has encountered in its inquiry several vulnerabilities within psychology research at Tilburg University, which lead to the following recommendations.

The relevant psychology-related scientific publications must give satisfactory specifications of how and where the data were collected.

Far more than is customary in psychology research practice, research replication must be made part of the basic instruments of the discipline. Research data that underlie psychology publications must be held on file for at least five years after publication, and be made available on request to other scientific practitioners. This rule is to apply not only to raw laboratory data, but also to completed questionnaires, audio and video recordings, etc. The publication must state where the raw data reside and how to access them.

The chronological distribution of the tasks made it inevitable that some people were interviewed by more than one committee.

Appendix 1 Levelt Committee interviewees

Dr Y.R. Avramova
Prof. I. van Beest
Dr L.M. van den Broek
Drs A.J. van Doorn
Prof. P. Eijlander
Drs M. Elshout
Drs E.R.K. Evers
Prof. G.L.M. van Heck
Dr F. van Horen
Dr Y. Inbar
Dr Janneke F. Jolij
Prof. G. Keren
Prof. E.J. Kraemer
Dr J. Lammers
Dr D. Lerouge
Drs L.A.J.G. van der Linde
Prof. S.M. Lindenberg
Drs M.H.C. Meijers
Drs M.H.J. Meijs
Dr R.M.A. Nelissen
Dr M.K. Noordewier
Prof. F.G.M. Pieters
Dr L. Renkema
Dr K.I. Ruijs
Prof. C.G. Rutte
Dr S.A. Schwinghammer
Prof. G.R. Semin
Prof. T.M.M. Verhallen
Prof. A.J.R. van de Vijver
Prof. A.J.J.M. Vingerhoets
Dr C.J. Wiekens
Drs J. van Wolferen
Prof. M. Zeelenberg

Appendix 2 Noort Committee interviewees

Prof. J.M.F. ten Berge
Prof. H. Blanton
Drs R. Boerrigter
Dr J. W. Bolderdijk
Dr A. van den Bos
Prof. A.P. Buunk
Prof. A. Dijksterhuis
Dr M.A.J. van Duijn
Prof. E.H. Gordijn
Dr J.D.M. Grob
Dr J.A.M. Heesink
Dr Janneke F. Jolij
Dr E. Kamans
Prof.. H.A.L. Kiers
Dr A. Klapwijk
Dr S.J. Ko
Dr W. Koomen
Dr J. Lammers
Prof. S.M. Lindenberg
Dr M. Maringer
Dr D. Marx
Dr K. Massar
Prof. D. Muller
Dr H.K.E. Oldenhuis
Prof. S. Otten
Prof. K.I. van Oudenhoven-van der Zee
Prof. J.L. Peschar
Dr W.J. Post
Prof. T.T. Postmes
Dr L. Renkema
Dr K.I. Ruijs
Dr S.A. Schwinghammer
Prof. R. Spears
Prof. J. Suls
Prof. A. Tesser
Dr D. Trampe
Dr S.W. van der Velde
Prof. E. van der Vliert
Dr C.J. Wiekens

Appendix 3 Mr Stapel's response to the draft interim report

October 28, 2011

Dear committee,

I have read the report with a sense of dismay and shame. Unfortunately my present state does not permit me to assess this report completely for any factual inaccuracies.

I must emphasize that the errors I have made were not motivated by self-interest. I do not identify with the picture that has been sketched of a man who has attempted to use young researchers for his own gain. I have made mistakes, but I was and remain sincerely committed to the field of social psychology, young researchers, and other colleagues. For this reason I regret the suffering I have caused others.

I intend to issue a written statement this coming Monday.

Diederik Stapel