

Research Review
Pedagogics and Education Science

Quality Assurance Netherlands Universities (QANU)
Catharijnesingel 56
PO Box 8035
3503 RA Utrecht
The Netherlands

Phone: +31 (0) 30 230 3100
Telefax: +31 (0) 30 230 3129
E-mail: info@qanu.nl
Internet: www.qanu.nl

Project number: Q404

© 2013 QANU

Text and numerical material from this publication may be reproduced in print, by photocopying or by any other means with the permission of QANU if the source is mentioned.

CONTENTS

Foreword by the committee chair	5
1. The review committee and review procedures	7
2. General remarks	15
Assessment at Institute and Programme level	25
3. Leiden University	27
4. University of Groningen	43
5. University of Twente	59
6. Utrecht University	69
7. University of Amsterdam	83
8. VU University Amsterdam	95
Appendices	111
Appendix A: Explanation of the SEP criteria and scores	113
Appendix B: Programme for the site visits	115
Appendix C: Short Curriculum Vitae of the committee members	121
Appendix D: Citation-analysis by Ad Prins	123

FOREWORD BY THE COMMITTEE CHAIR

The purpose of this report is to present a reliable picture and evaluation of the research in pedagogics and educational sciences at Leiden University (LEI), University of Groningen (RUG), University of Twente (UT), Utrecht University (UU), VU University Amsterdam (VU) and University of Amsterdam (UvA) conducted in the period 2006-2011, and to give feedback on the research management and quality assurance.

This report follows the Standard Evaluation Protocol 2009-2015 for Public Research Organisations (SEP) that was developed by VSNU, KNAW and NWO. The review committee, which was composed of scholars from various nationalities and different pedagogical and educational subfields, was supported by QANU (Quality Assurance Netherlands Universities).

As chairperson of this committee, I would like to thank the official representatives and other staff of the institutes and programmes under review for their loyal cooperation in all phases of this assessment process. I also wish to thank the members of the review committee for their willingness to participate in this assessment and for the dedication with which they carried out this delicate task. Finally, I wish to thank the Quality Assurance Netherlands Universities, and in particular Dr. Floor Meijer and her co-secretary Drs. Mariëlle Klerks, for their great support in the organisation of this assessment and in the production of the assessment report, but also for their inexhaustible efforts to make the committee members feel at home in Utrecht.

Lieven Verschaffel
Chair of the Committee

1. THE REVIEW COMMITTEE AND REVIEW PROCEDURES

Scope of the assessment

The Pedagogics Committee was asked to perform an assessment of the research in pedagogics and education science at Leiden University (LEI), University of Groningen (RUG), University of Twente (UT), Utrecht University (UU), VU University Amsterdam (VU) and University of Amsterdam (UvA). This assessment covers the research conducted in the period 2006-2011.

In accordance with the Standard Evaluation Protocol 2009-2015 for Research Assessment in the Netherlands (SEP), the committee's tasks were to assess the quality of the institutes and the research programmes on the basis of the information provided by the institutes and interviews with the management and the research leaders, and to advise on how it might be improved.

Composition of the committee

The composition of the Pedagogics committee was as follows:

- Prof. L. (Lieven) Verschaffel (chair), Professor at the KU Leuven, Belgium;
- Prof. M. (Mary) Dozier, Amy E. DuPont Chair of Child Development, Professor of Psychology, University of Delaware, USA;
- Prof. M.T. (Mark) Greenberg, Edna Peterson Bennett Endowed Chair in Prevention Research, Professor of Human Development and Psychology, The Pennsylvania State University, USA;
- Prof. E. (Eckhard) Klieme, Director of the Center for Research on Educational Quality and Evaluation, Johann Wolfgang Goethe University, Frankfurt am Main, Germany;
- Prof. H. (Heikki) Lyytinen, Professor at University of Jyväskylä, Finland;
- Prof. N. (Neil) Mercer, Professor of Education at Cambridge University, UK;
- Prof. K. (Karine) Verschueren, Professor at the KU Leuven, Belgium.

Dr. F. (Floor) Meijer of QANU (Quality Assurance Netherlands Universities) was appointed secretary to the committee. A short curriculum vitae of the committee members can be found in Appendix C.

Independence

All members of the committee signed a statement of independence to guarantee that they would assess the quality of the institutes and research programmes in an unbiased and independent way. Any existing personal or professional relationships between committee members and the programmes under review were reported and discussed in the committee meeting.

Data provided to the committee

The committee received detailed documentation consisting of the following parts:

- Self-evaluation reports of the units under review, including all the information required by the Standard Evaluation Protocol (SEP), with appendices;
- Key publications per research programme, with a maximum of five articles/books;
- Publication lists of staff members per research programme;

- Results of the bibliometric study carried out by Ad Prins.

Committee remarks regarding the data provided

The committee appreciates the efforts of the institute and programme leaders to provide all the requested information. However, it did encounter some problems with the accuracy, completeness and standardisation of this information, particularly the quantitative information. In some self-evaluation reports the attached lists of publications did not differentiate between internal and external authors and/or did not mention the impact factors of the journals. The commission asked those institutes to send a revised list. Other unclear aspects or inconsistencies in the reporting of the research output (Standard Evaluation Protocol [SEP], Table 5.3) related to definition issues ('internal' vs. 'external' PhD theses, 'professional' and 'popular' publications). With respect to grant-earning capacity, the committee found it unfortunate that most institutes/programmes did not provide an exhaustive list of the direct and indirect grants obtained per year. Depending on the nature and seriousness of these queries, the committee decided to resolve them by email before or after the visit or through questions during the visit.

The committee further notes that most, but not all, institutes followed the instructions for preparing the self-evaluation reports. Even so, there were considerable differences in the level of detail. For example, variations were observed in the description of the research infrastructure, the technical facilities, and the annual distribution of the three basic kinds of funding over the various programmes. These differences were particularly large with respect to the information provided in the attachments, dealing with elements such as academic reputation, awards and grants, indications of societal relevance, and 'quantifications of research quality' such as H-indices (based on Google Scholar or Web of Science) of staff members. Variations were also observed in the information provided on the number of research fte for the individual institute and programme members over the years of the assessment (and on precisely how these research fte's were determined). Given that the SEP does not provide clear instructions on these matters and that the committee did not want to increase the administrative burden further and endorse the trend towards even more quantification of research assessments, the committee has decided not to ask for additional information. Information (on H-indices or other quantitative productivity and/or impact measures and/or societal relevance) that was spontaneously provided by the institute/programme or collected by individual committee members has been used in the assessment. But the absence of such information in a self-evaluation report was never used against an institute or programme.

Procedures followed by the committee

The committee proceeded according to the SEP 2009-2015. Each programme was assigned to two reviewers, who independently formulated a preliminary assessment. The first reviewer was chosen on the basis of his or her expertise in the domain of the programme; the second reviewer was chosen to provide a more general, complementary perspective.

Before conducting interviews with representatives of the institutes and programmes under assessment, the committee was briefed by QANU about research assessment according to SEP and discussed the preliminary assessments. It also agreed upon procedural matters and aspects of the assessment. For each university it discussed the self-evaluation report, key publications and the preliminary findings of all research programmes and the institute before starting on the interviews. The first reviewers led the interviews, and then the second reviewers and the other committee members were given the opportunity to ask questions. After each interview the committee took some time to prepare a preliminary assessment.

After concluding the interviews at each university, it discussed the scores and comments of the institute and programmes.

The interviews took place on 24, 25, 26 and 27 March 2013 (see the schedule in Appendix B) at a central location in Utrecht, the Netherlands. The site visit consisted of 60-minute interviews with the management of the research institutes and 45-minute interviews with (1) the leaders of each of the research programmes and (2) a selection of PhD students.

Due to illness, Prof. Klieme was not able to attend the site visit.

At the end of its meeting in Utrecht, the committee discussed the scores and comments of all 14 programmes and 6 institutes. The final assessments are based on the documentation provided by the institutes, the key publications, and the interviews with the management and the leaders of the programmes. The texts for the committee report were finalised through email exchanges. The first assessor was responsible for writing the draft assessment and for sending it to the second assessor for amendment and/or approval. After it had been approved by both assessors, the assessment was inserted into the report. After receiving all assessments, the secretary compiled the report and returned it to the committee for a final approval. The approved version of the report was presented to the faculties for factual corrections and comments. The final report was presented to the Boards of the participating universities and was printed after their formal acceptance.

The SEP 2009-2015 uses a 5-point rating scale (see Appendix A). The committee slightly adapted this rating scale as it quickly became clear that the pedagogical and educational research in the Netherlands is generally of a good to very good level, implying that most ratings involved the higher end of the scale. To allow differentiation in this rather narrow range, the committee decided to extend the 5-point scale to a 9-point scale (1, 1.5, 2, ..., 4.5, 5). The .5 was used to indicate that a programme is between two integer ratings. The committee wants to emphasise that it has taken very seriously the SEP request to consider the full range of this five point scale and to apply the scores according to the descriptions given in the SEP. This means, for instance, that a score of 5, which expresses 'world class' can, by definition, only be given very rarely, and that a score of 3,5 refers to research that is better than good and almost very good, not only according to national but also to international standards, which have, over the past years, become increasingly high. In doing so, the committee has tried to resist the ongoing trend of inflation of scores, which leads to a situation wherein variation between scores disappears and scores become meaningless (Cf. Rathenau Instituut, 'Twenty years of research evaluation', July 2013). The committee insists that the institute and programme leaders, as well as others who may be interested in and make use of these scores, should interpret them accordingly, and, moreover, always see them in relation to the qualitative comments.

Like the previous assessment committee (cf. 'Research Assessment Pedagogics and Education Science 2007', RAPES), the committee decided that it would not attempt to give an overview and discussion of the state-of-the-art in pedagogics and educational sciences in the Netherlands from an international perspective, but rather restrict its task to the evaluation of the institutes and programmes involved in the assessment. It took that decision, first, because it was not feasible to draw up a complete overview of the position of Dutch pedagogical and educational research in the international scene; and, second, because the participating universities and programmes are only part of the national scene in these areas of research. Indeed, some institutes are missing, and in the participating institutes, not all programmes related to pedagogics and educational research were involved in the assessment.

In making its assessment, the committee followed the SEP instructions concerning the four main criteria (and their respective sub-criteria), even though it sometimes struggled with a clear demarcation between these four distinct categories. In particular, the distinctions between quality and productivity, and between productivity and societal relevance, elicited quite some concern and discussion. But the committee has nevertheless tried to work with the definitions and operationalisations of these four criteria as outlined in the SEP as far as possible. How the committee dealt with these matters for the four main criteria is specified below.

Quality

In line with the SEP instructions, quantitative as well as qualitative information about the prominence, prestige, and impact of an institute's or programme's scientific output have been used to assess its quality. As far as academic publications are concerned, this means that qualitative valuations of the prominence or prestige of the journal(s) (essentially international) and/or other kinds of academic publications were taken into account, as well as quantitative measures that aim to indicate the quality and/or impact of a researcher's or programme's scientific products. If available, the following quantitative elements were used in the assessment (albeit not necessarily all of them or all to the same extent):

- Number of items in Web of Science (particularly for the period 2006-2011);
- Number of items in Google Scholar (particularly for the period 2006-2011), taking into account that for certain subdomains of pedagogy and educational sciences, the Web of Science may be too narrow and/or too biased (cf. the section about the bibliometric report by Prins);
- Number of items in top and subtop journals in Web of Science (somewhat pragmatically, the term 'top journals' was restricted to publications that are in the top 10% of a given Web of Science category, while the term subtop was used for journals that are among the highest 25% of a certain category);
- Number of citations of Web of Science articles (particularly for the period 2006-2011) – both individual articles and the total of articles – as a further indication of an institute's or programme's impact;
- H-indices of tenured and possibly non-tenured staff (particularly for the period 2006-2011), as a relevant combined measure of scientific impact and productivity.

For certain programmes (e.g., ones that operate at the frontier of social sciences and arts and, thus, within different publication and citation cultures, and/or that operate within smaller scientific niches), the above-mentioned quantitative indicators of quality and impact were considered less applicable. Quantitative indicators of the quality of a programme's research output were consequently complemented and balanced with qualitative appreciations based on the committee members' prior knowledge of the programmes and their outcomes, as well as a careful reading of the key publications and publication lists.

The following qualitative indicators of academic reputation were used in the assessment: editorships and membership of editorial boards of journals and book series, responsibilities in scientific organisations and evaluation committees, invitations to contribute to prestigious books, invitations to give plenary lectures at scientific conferences, memberships of scientific advisory committees, and various kinds of honours and prizes, etc.

In line with the SEP instructions, funding strategies and capacities have been included in this criterion too. The committee looked in particular at the following elements:

- (Changes in) the absolute amount of direct and non-direct funding during the review period;
- (Changes in) the amount of funding gained from research grants and/or contract research as compared with the amount of direct funding;
- A qualitative assessment of the nature (prestige, variety, international...) of the indirect funding obtained, with particular attention paid to grants obtained from the EC and other international funding agencies.

To allow for some appreciation of a programme's funding success in relation to its size, the available quantitative information about an institute's or programme's funding success (the development in it) was evaluated not only in absolute terms, but also in relation to the size of the programme team; and, more particularly, in terms of the total number of research full-time equivalent (fte) for tenured staff for the whole assessment period. The committee therefore used the information provided in the self-evaluation report for every programme (i.e., Table 5.4 of the SEP) to compute the amount of indirect funds (i.e., research grants and contract research) obtained over the six years of the assessment and divided it by the total research fte for tenured staff. The committee is aware of the pitfalls involved in using this measure and was therefore very cautious in using it for comparative purposes, especially for comparisons between institutes.

Finally, with respect to the quality of the research, the committee also looked at the quality of the PhD training and supervision, although this assessment was made at the level of the institute rather than at the programme level.

Productivity

With respect to productivity, the committee noted that, according to the SEP, it should (1) include both the productivity strategy and the effective productivity, (2) refer to professional publications and output for a wider public and not just to scientific publications and PhD theses, (3) judge the output in relation to the institute's mission and resources, and (4) maintain a good balance between quantitative and qualitative elements, with any quantification using international standards as far as possible and where relevant.

To arrive at the productivity score, the committee began by considering the quantitative data provided on scientific publications and PhD theses, and also on professional publications and products, as listed in Table 5.3 of the self-evaluation reports. It looked at the spread of the output over the different output categories, its evolution over the six years of the evaluation period, and the output in relation to the available fte. In order to make systematic comparisons across the programme reports, the following extra measures were always computed and reported:

- Total scientific output / total fte research staff;
- Total refereed English (and other non-Dutch languages) journals / total fte research staff;
- Total professional and popular publications / total fte research staff;
- Total PhD / total fte tenured staff.

However, depending on the nature and mission of the programme and/or on the composition of the programme's research staff, the committee also included (and sometimes even gave priority to) other quantitative indicators, such as the number of international books and/or book chapters, looked at the tenured staff rather than the total research staff to value a programme's productivity relative to its research capacity, and/or took into account the

development in the research output during the assessment period. To determine the ultimate productivity score, quantitative measures given in or derived from the self-evaluation reports were always combined with qualitative assessments based on the SEP principles. In sum, a programme's productivity score does not follow 'automatically' from the quantitative indicators of productivity (in relation to available fte) given in Table 5.3 of the self-assessment reports.

Societal relevance

With respect to societal relevance, the committee looked for and evaluated the societal quality of the work, its potential and actual societal impact, and its 'valorisation' (as outlined in the SEP), based on the available evidence in the self-assessment reports and their attachments, and additional information provided during the meetings with the institute and/or programme leaders. In some cases, extra information was sought on the internet.

Viability

This criterion regards the institute's or programme's ability to reflect upon and react to important changes in its environment. It involves issues such as concentration and integration of research lines, collaboration with other institutes/programmes, strategic planning, etc. (see SEP). To assess a programme's viability, the committee sometimes also took into account spontaneously provided information on the most recent developments (which occurred after the end of the assessment period) provided in the self-assessment reports and/or during the meetings with the institute and/or programme leaders.

Position of the committee towards the Prins report

The committee took note of the carefully prepared report 'Bibliometric Analysis of Research Programmes in Education and Pedagogical Sciences' by A. Prins (cf. appendix D), which had been ordered by the directors of the different institutes for the present assessment. This bibliometric analysis provides an international comparison for the research programmes involved in the present assessment with respect to the following question: 'How does the impact of the Dutch educational and pedagogical scientists compare to the impact of their international colleagues?' The analysis is based on a selective subset of the output of these programmes. The programmes were asked to send in ten publications for each year between 2006 and 2011, resulting in a total of sixty publications per programme. (In one case only thirty publications were involved, and one programme decided not to participate in the analysis.)

The analysis was designed to address the difficult task of international comparison, while also taking account of disciplinary differences between programmes in terms of publication and citation culture. Because the programmes' output is made up of a wide range of journals and other publication formats (including books and chapters in books, conference proceedings, scientific reports, etc.), Google Scholar was chosen as a more suitable data source (rather than e.g. Web of Science). While the results of the analysis are presented per programme, it must be emphasised that the results are not suited for comparing the programmes.

To address the issues of both international comparison and disciplinary diversity, Prins set up three comparisons that reflect various ways to define the disciplinary domains of the fields. For each domain he computed quotients that represent the international average impact of that domain. These quotients are based on totals of citations after publication divided by the number of documents published (data for impact quotients and domains were derived from Scopus). These data were then compared with the impact of each programme, which was

computed as the total of citations of the publications for that programme divided by the number of publications involved. The three comparisons are:

- **Group-Specific Domains:** The basis of comparison in this approach is formed by the six journals in which each programme publishes most often. Averages of the citations of articles of programmes published in these six journals are compared with the average impacts of these journals.
- **Discipline-Specific Domains:** The average impact of each programme is compared with the average impact of journals in two domains that relate to the disciplinary differences in the field: Education and Pedagogical Sciences. The journals are part of a long list of 457 journals taken from several Scopus science categories relevant to these domains. Each journal in the list has been categorised by researchers in the field as representing the domain of either Education (322 journals) or Pedagogical Sciences (118), or both (17). As an indication of the relevance of the two domains for each programme, a percentage is given of how many of its submitted publications are in journals of either domain.
- **General Scopus Domains:** The impact of each programme is compared with the impact of journals in three general Scopus domains. The relevant Scopus domains are: Education, Psychology Miscellaneous, and Psychiatry & Mental Health. As an indication of the relevance of the Scopus science category for each programme, a percentage is given of how many of its submitted publications are in journals of this category.

Of course, because programmes differ considerably in size, the selection base within programmes depends on the size of the programme. The number of highly visible publications was more limited in the smaller programmes and may therefore have resulted in lower impact averages. For that reason, it is emphasised that the results of the analyses are not suited for comparison of the programmes, but may allow a comparison of the impact of selected publications of the programmes against international standards.

The committee noted that one programme strongly argued against the validity of the citation analysis based on Google Scholar in its self-assessment report; to better demonstrate its excellence, it conducted an additional analysis of the same data set based on the Web of Science criteria which it considers to be stricter and more accurate. One other programme, belonging to another institute, decided to withdraw from the analysis because it disagreed with the way in which the inclusion criteria for certain types of publications were operationalised. Furthermore, during the meetings with the institute and programme leaders, some additional critical comments about this analysis and its outcomes were made. On the other hand, other institutes supported Prins's endeavour, expressing their appreciation for the attempt to design measures that are tuned to the research field in general and specific subfields in particular, and arguing that the pleas in favour of the Web of Science as the critical data base are largely unjustified, especially for assessing productivity and impact in the fields of pedagogy and educational sciences.

Taking everything into account, the committee decided to make use of Prins's report, particularly in the general part of the assessment, as *one* indicator of the quality and impact of Dutch research in pedagogy and educational sciences in an international perspective. The results of this report were also used in the individual programme assessments, though always along with other quality indicators and taking into account the above comments of Prins's approach.

Previous assessment reports

The committee considered it essential to look at the latest research reviews (final and/or mid-term) per institute and per programme. It thanks all of the institutes for helping to identify the most relevant previous assessment reports, for making these reports available, and for providing a helpful clarifying commentary in some cases. These previous reports did not directly affect the qualitative evaluations and the accompanying scores. Rather, they represented an additional element that was taken into account in the overall assessment. In this respect, it is important to note that several self-assessment reports contained explicit references to, and comments on, previous self-assessments.

2. GENERAL REMARKS

Introduction

The committee's overall impression was that the reports and presentations of all institutes and programmes revealed serious and mostly successful attempts to further improve the quality, productivity, societal relevance and viability of their research. They acted upon the conclusions and recommendations of the last evaluation and adapted to developments within the relevant discipline(s). They also took account of recent, rapid and profound changes to the research policy at the university, national and international levels. The committee further noted that steps had been taken to re-focus existing programmes; to integrate previously distinct programmes into one larger programme; to install completely new programmes in strategically important research areas (such as educational neuroscience); to improve the internal organisation of the research at the institute and/or programme level; to refine and enhance the criteria for the selection and promotion of staff (especially tenured track); to renew and enrich the technical equipment and research infrastructure; to improve the intensity, efficiency and quality of the doctoral training and supervision of PhD students and to establish or improve research masters that act as stepping-stones to research within the programme(s); to further enhance the quantity and quality of academic and professional output; to substantially increase the amount of non-direct funding, with special attention to prestigious (inter)national funding categories; to strengthen strategic positions within newly established interdisciplinary entities (i.e., research institutes, networks, or areas); and to establish or strengthen links with other institutes and programmes (inter)nationally.

Various paradigms within pedagogy and education research

Since the last RAPES-2007 report, the process of 'psychologisation' of the pedagogical and educational institutes and programmes has continued, with accompanying changes in the dominant theoretical frameworks, methodological approaches, and publication practices and cultures. This ongoing movement seems to have been enhanced by the remarkable increase in the inclusion of concepts and methodologies from the fields of cognitive neuroscience and biogenetics. All but one institute involved in the current assessment has now – at least to some extent – incorporated these neuroscientific and/or biogenetical perspectives and methodologies into their programmes, or have set up completely new ones that capitalise on these new perspectives. A first important consequence of this trend is the increased interdisciplinarity of the research programmes and the intensified cooperation with scholars and teams from other, particularly bio-medically oriented, disciplines (see further). Another consequence is that pedagogical and educational research is becoming far more expensive than it used to be. Arguably, this may jeopardise the future of this kind of research, but it may also threaten, in an indirect way, the continued funding of the other, more traditional, types of pedagogical and educational research.

Complementary to the 'psychologisation' of pedagogical and educational research, the dynamics in the field may, according to a response from one institute to the committee's draft report, also be described as 'pedagogisation' of important parts of psychology, in the sense that psychologists have felt increasingly attracted to, and have become more and more involved in, the field of pedagogy and education. In other words, the general mission and concrete themes of the discipline of pedagogics and educational science have led an increasing number of Dutch (cognitive, developmental and social) psychologists to invest their intellectual capital in this discipline, rather than continuing to pursue a strictly psychological perspective. This has not only helped them to broaden their theoretical perspective and methodological repertoire; it has also helped them to find a better balance between more fundamental and applied research, by choosing research topics that increase

understanding of human learning and development, while at the same time providing building blocks for its improvement, and also by investigating these topics in more complex and ecologically valid settings.

Parallel with the previously described developments, programmes that follow philosophically and/or historically and/or culturally oriented approaches to pedagogy and education rather than the empirical c.q. psychology-oriented approach seem to have suffered from a further downsizing. Some of these programmes were forced to merge with other (smaller, threatened), more empirically oriented programmes (with senior staff members with theoretical interests having to find their place within bigger, more empirically oriented pedagogical or educational programmes), while those programmes that remained on their own seem to have a hard time surviving. In response to external factors and for reasons related to intrinsic changes within their subdisciplines, these programmes have searched for new ways to complement their traditionally dominant philosophical and/or historical perspectives with those from sociology, cultural studies, activity theory, etc., to connect their areas of research activity more closely to contemporary themes (e.g. diversity), and to consider educational and societal reality with a more 'experimental' attitude. However, due to their quite different paradigms and accompanying research and publication cultures, these (sub)programmes continue to experience difficulty in meeting the institutional targets for impact and productivity that dominate in the social and behavioral sciences, including pedagogy and educational research. Nevertheless, for reasons related to all three main areas of a university's mission (i.e., research, teaching, and societal service), there is still a need for high-quality research that does not follow the dominant empiricist c.q. (bio-)psychologically oriented research paradigm, but considers pedagogy and education from another, broader, socio-cultural perspective, with its own accompanying research methods. One question that arises is whether this research is ideally situated in separate (smaller) research programmes, or if it would be more productive to integrate it within larger pedagogical or educational programmes.

Multi-/inter-/transdisciplinarity

Another major observation concerns the institutes' and programmes' movement towards multi-/inter-/transdisciplinarity (hereafter: interdisciplinarity), although it should be acknowledged that for many institutes and programmes, this development already started before this assessment period.

All self-evaluation reports endorse – rightly, in the committee's view – the position that research in pedagogics and educational sciences has become increasingly interdisciplinary. The reports explain and document what steps have been taken in that direction, in terms of both research policy and actual research. Probably the most remarkable trend in this respect is the intensification of the collaboration with scholars from the biomedical sciences, but researchers are also involved in establishing new research lines with other disciplines like administration, technology, cultural studies, etc. This is evidenced, among other things, in the increasing breadth of the journals in which pedagogical and educational researchers publish.

This tendency towards interdisciplinarity cannot be separated from the installation of large-scale, multidisciplinary entities (research institutes, networks, areas) within the universities. Between 2006 and 2011, all universities involved in the current assessment took the initiative to get more involved in the 'profiling' of their research around major sectors or themes with high societal relevance such as energy, food, life sciences, high tech, etc. The general idea was that this should result in large, strong, interdisciplinary research entities that would be well prepared for the increasingly selective national and especially international competition for

research funds. In some universities these interdisciplinary entities took the form of new institutes that replaced existing institutes or programmes, while others established interdisciplinary research priority areas while maintaining the original programmes or while only a selection of researchers from these programmes collaborate in the new profiling programmes. These new entities are typically built around interdisciplinary top priority areas ('topsectoren'), involve the university's best research teams, and – last but not least – attract an increasingly large part of the available direct funds. In some cases the pedagogical and educational institutes or programmes succeeded in obtaining a central position in those interdisciplinary entities, while in other cases they are only involved as smaller and secondary partners, or are still trying to find their place in such a network.

The committee applauds the creation of these larger, strong, interdisciplinary research networks that may help to better prepare institutes and programmes for the demands of interdisciplinary research, for improved training of the next generation of scientists, for addressing current societal and educational problems, and for the increasing competition for research funds. With respect to the question of what would be the institute's and programme's preferred position in these large-scale networks, the committee recommends, in general, to strive for an active and prominent involvement. At the same time, it is important to also preserve, where essential, the embedding of the research institutes in discipline-based faculties that have traditionally housed the domains of pedagogical and educational research. It is the committee's conviction that a strong disciplinary research identity and research tradition are prerequisites for fruitful interdisciplinary collaboration and for maintaining prominent positions in these interdisciplinary initiatives.

Hereafter, the committee summarises some general observations with respect to the four criteria of the assessment. But, first, it presents some general observations about the mission and scope of the research.

Mission and scope of Dutch pedagogical and educational research

In terms of mission, all institutes and programmes aim to increase the scientific knowledge base for pedagogy and education and to contribute to the solution of problems related to the pedagogical and education practice. There are substantial differences, however, between institutes and programmes with respect to the balance between the two aspects.

With their research, the more pedagogically oriented programmes cover themes such as: typical and atypical development of children and adolescents; characteristics and determinants of cognitive and socio-emotional development in children, adolescents, and (young) adults; assessment, prevention, and treatment of psychosocial and learning/cognitive risks and problems in these various age groups; the interaction of biological and genetic factors with proximal and distal environmental factors to understand how both biology and environmental affordances contribute to developmental outcomes; severe behavioural and learning problems; disabilities; parenting and childrearing in general and in challenging situations in particular; child and youth care.

Recurrent themes of the educational research are: effects of learner characteristics, educational interventions, and school and contextual factors on preschool and school learning and learning at the workplace; diagnosis and education of children and other people with various kinds of disabilities; classroom interaction; teaching and learning in specific curricular domains; teaching and teacher education; educational technology and educational assessment; educational innovation and school leadership. Most educational programmes also involve intervention research. Ample attention is paid to cognitive and metacognitive processes and

outcomes, but there is also much interest in social and affective aspects of learning, and social issues of segregation, selection mechanisms, and school dropout.

While there are some differences with respect to the programmes' dominant theoretical lenses and preferred research methods, pronounced paradigmatic tensions seem rare. The common theoretical background of the programmes in pedagogy and child studies is made up of bio-ecological models of development that address interactions between various factors at different levels of analysis. Likewise, educational programmes look at factors at the micro-, meso-, and macro-level that affect the nature and quality of learning and instruction. At the micro-level, cognitive/rationalist and situative/pragmatist/sociohistoric views on knowing, learning and teaching predominate. Thus, there are many points of convergence between the more pedagogically and educationally oriented programmes and, since that last review, very positive actions to successfully integrate concepts and research programmes across the two broad areas.

Quality

The main conclusion that can be drawn from the self-evaluation reports, the enclosed documents, and the meetings with the managers, leaders and PhD students of the institutes and programmes is that Dutch research in pedagogy and education is generally of (very) high quality and has further improved since the previous assessment. Evidently, the programmes differ in quality, as evidenced in the programme assessment reports and scores, but in almost all cases the quality was considered to range between very good and excellent, and the few relatively weaker scores were due to certain specific aspects or parts of a programme. Most research teams that performed very well during the previous assessment period succeeded in maintaining their quality, while others managed to augment it. Of course, when giving scores, the committee has taken into account that, since the previous assessment, the international research scene has also changed substantially with respect to pedagogy and education. Therefore, the criteria for international excellence have become much stricter.

This positive general conclusion with respect to the quality of pedagogical and educational research in the Netherlands is based on the fact that all programmes publish regularly in the leading journals of their field. Most programmes report at least several, if not multiple, publications in the top or subtop journals of their subfield(s) of psychology and educational sciences and sometimes also in the most prestigious transdisciplinary outlets such as *Science*. Moreover, during the review period many programme leaders and/or other programme members were invited to write chapters in/for the most prestigious handbooks, encyclopedias or other specialist books in their fields, or have themselves edited or written books published in leading international book series or by leading international publishers.

This positive conclusion is further supported by Prins's bibliometrical study. This citation analysis has shown that the impact of the Dutch programmes as a whole is much larger than the average impact of the leading journals in the field, and that only in a few exceptional cases a programme did achieve somewhat less impact than would be expected given the various standards used in the analysis. Even though one could raise concerns about its reliability and validity (particularly because of its exclusive reliance on Google Scholar), the analysis nevertheless provides additional support to the above positive conclusion.

Quality is also underlined by the fact that many individual programme leaders enjoy an excellent international reputation. They received many signs of recognition in the form of national and international academic honours and awards, editorships and memberships of editorial boards of numerous international journals (including several top and subtop journals

in the fields of psychology and educational sciences), invitations for major keynote lectures at leading international conferences, for memberships on PhD juries abroad, and for various kinds of scientific evaluation, programming, or advisory committees in other European countries and worldwide. Clearly, Dutch scientists are prominent participants in the international scene of pedagogical and educational research.

Another quality indicator is their capacity to obtain grants. Table 1 presents the total amount of funding and the percentage per year and funding category across the six institutes. As shown in this table, the total amount of funding increased further during the assessment period. As will become clear in the next part of the assessment report, most institutes and programmes also increased their funding from year to year during the current assessment period, although this was not always the case. Table 1 reveals that the increase in the total amount of funding is essentially due to the impressive increase in the category of research grants. As a result, for the whole assessment period, about half of the funding is direct, while the two other categories represent about a quarter of the total funding.

Year	Funding category			Total
	Direct	Grants	Contracts & Other	
2006	10710 (51%)	4280 (20%)	6133 (29%)	21040
2007	11219 (51%)	4601 (21%)	6309 (28%)	22152
2008	11974 (47%)	4674 (18%)	8815 (35%)	25266
2009	12187 (51%)	5958 (25%)	5696 (24%)	23840
2010	12598 (50%)	6915 (27%)	5876 (23%)	25391
2011	12589 (43%)	9852 (34%)	6647 (23%)	29020
Total	71277 (48%)	36279 (25%)	39476 (27%)	146711

Table 1: Funding and percentage per year (between parentheses) and funding category, across the six institutes (x1000 euros).

There are also considerable differences between the institutes and programmes with respect to how the funds are distributed over the three categories and how this distribution has evolved over time. However, it should be noted that this might, to some extent, be due to differences in the definition and operationalization of the distinct funding categories in the various self-assessment reports.

As emphasised in most of the self-assessment reports, it is to be expected that the total amount of money for which these institutes and programmes can compete within their faculty, university, and country will most probably not rise in the coming years but rather decrease, partly as a result of the Dutch government's and the universities' strong tendency to invest the available money into a small number of top priority areas that do not always fit optimally with the domains of pedagogy and education. It seems that institutes and programmes will have to look for new opportunities for funding at the international level (especially the European level). Although most programmes report some successful steps in this direction, with the exception of a couple of successful programmes, the quantity of such international grants is generally still relatively small. Some institutes reportedly offer strategic and technical support to researchers who develop initiatives in this direction, which is, of course, to be welcomed. It can be expected that, with the creation of larger programmes within the institutes and of the interdisciplinary research networks within the universities, the initiatives and chances of success will significantly increase. Striving harder for EC grants seems one of the most common strategic intentions for the future for all institutes and programmes.

As a final, more general reflection, the committee would like to comment that augmenting the amount of grant money (and the amount of non-tenured and tenured staff that goes with it) should not be considered a goal in itself but rather as a means to accomplish a programme's scientific mission.

Productivity

Productivity is generally (very) high. Taken as a whole, Dutch researchers in pedagogy and education produce a very large number of academic publications in the form of international journal articles as well as international books and book chapters. Institutes and programmes differ quite a lot, though, with respect to the attention they want to pay to Dutch academic publications and to professional publications and products. According to the committee, national language publications remain important for keeping the scientific study of national educational practices alive and for serving the professional and public audience.

Particularly with respect to academic publications, one can raise the question of whether productivity is given too much weight and is sometimes valued at the expense of the quality of the research. In this respect it was interesting to notice, both in the written self-assessments and in the meetings with institute and programme leaders, how the possible trade-off between quantity and quality is becoming an issue of even more serious concern and debate, and how the research managers are starting to think of ways to give a more prominent role to quality rather than sheer productivity in the allocation of the available direct research funds and/or in selection and promotion of staff. A possible (radical) step in that direction could be to simply stop asking for numbers and complete lists of the academic and professional output being produced by an institute or programme, and focus on a small part of it (e.g., a programme's ten best publications per year, or each researcher's five or ten key publications for the whole assessment period), which would then be subject primarily to a more qualitative review.

In line with the SEP instructions, the committee relied heavily (but not exclusively) on the information about productivity as provided in the self-assessment reports (and particularly in Table 5.3 of the SEP) to arrive at the productivity scores because, after all, productivity (as defined and operationalised in the SEP instructions) has been a major point of attention of all institutes and programmes during the whole assessment period, and because productivity will probably continue to play a major role in the academic competition among institutes,

programmes and individuals, in national and international grant acquisitions, etc. in the coming years. At the same time, the committee wants to argue that the current pressure to be productive has – in most if not in all cases – reached a point where it may no longer contribute to the quality of the research, the well-being of the researchers, and the betterment of society.

Societal relevance

The societal relevance and impact of the research and the researchers in pedagogical and educational sciences are generally high, which may not come as a surprise as domains like pedagogy and education study societally relevant topics such as youth and parental problems, children with learning disabilities, people with severe handicaps, education of immigrant children, teachers' professional development, computers in education, etc.

All Dutch institutes in this field have a long-standing tradition of spending a substantial part of their time and effort in doing research and development work with practical relevance to pressing societal issues and problems of everyday life, in disseminating the results of their studies to professionals and the wider public, and in contributing in numerous other ways to the betterment of society. Many of them report various signs that these contributions actually positively affect society and are appreciated, e.g., by making extensive use of research-based 'products', by making available grants for contract research projects and paying for endowed professorships, by giving the researchers a forum in popular media, by asking them to join various kinds of advisory committees, by giving them awards for their societal contribution, etc.

Compared to the scientific part of the institutes' and programmes' mission, there seems to be much less control over and rewards for the work and investments staff members do in the societal domain. For instance, only one of the universities involved in this assessment applies a system for the allocation of its available direct funds that also takes into account the amount of contract research funding acquired by an institute or programme. The other institutes and programmes seem to consider it to be an evident historical and current part of their academic mission.

In this respect the committee refers to the recent report of the Royal Netherlands Academy of Arts and Sciences (KNAW, 2013) entitled *Naar een raamwerk voor de kwaliteitsbeoordeling van sociaalwetenschappelijk onderzoek* that contains a profound discussion of the societal relevance of research in the social sciences as well as an excellent framework for a better categorisation and valuation of the societal relevance of this research. While the committee strongly supports this endeavour to search for a better balance between (the assessment of) the scientific and the societal value of research, in line with this KNAW report, it is convinced that the scientific quality of research remains a priority for university-based research institutes and programmes.

Viability

With only a few exceptions, the committee is very confident about the viability of the Dutch institutes and programmes in the domains of pedagogy and educational sciences. Several institutes and programmes participate in a number of university-wide multidisciplinary institutes, which will become increasingly important for their viability. The self-evaluation reports describe the multidisciplinary embedding as a strength and opportunity. It is thought to provide a favourable institutional setting. Interestingly, some institutes and programmes have themselves taken the lead in some of those strategically important interdisciplinary

institutes or other networking structures within or between Dutch universities. Some specific elements concerning viability (organisational issues, personnel issues) will be discussed below.

Organisational issues

In the previous assessment report (RAPES-2007), a general concern was expressed about the absence of a transparent and efficient research management system/structure in most universities/faculties. The previous report therefore pleads for streamlining and professionalising of administrative issues above the research programme level. The nature of the information provided in the current review, and the short available time for the meetings with the institute and programme management, did not always allow for an in-depth review of the complex and multi-layered organisation of research and research policy. Moreover, with the emergence of and participation in these higher-order interdisciplinary entities, the research management system has become even more multi-layered and complex. Especially in those cases where different members of one programme are participating in different interdisciplinary research teams, problems of internal coherence may arise.

Personnel issues

It is generally accepted that an optimal age structure of a research unit is characterised by a good distribution of the personnel at the various levels of an academic career, including post-docs and young faculty members. However, in several programmes, the actual age structure seems to be far from optimal, either because there are too few (senior) tenured staff for the proper guidance of the numerous postdocs and PhD students (especially in newly established programmes or programmes that have witnessed a quick and drastic growth in non-direct research funding), or because the relative shortage of non-direct funding resulted in a disproportionate share of senior tenured staff members who are close to retirement.

In several cases, the self-evaluation reports referred to the fact that one or more programme leaders had left the programme during the past assessment period or will do so soon. The committee found it remarkable that in many SWOT analyses, finding a successor of the same calibre was mentioned as one of the major threats for the future. It wants to underscore this fear, given that in many (but not all) programmes there seems to be a remarkably big gap between the international prominence of the full professor(s) and the associate and assistant professors. Remarkably, in several programmes this succession problem has been/will be temporarily resolved by allowing these departing programme leaders to stay active in the programme.

PhD training

Two types of PhD students are distinguished. 'Internal' PhD students work within the universities and are financed directly by them or by grants from NWO and the EU, or by contracts with others. 'External' PhD students ('buitenpromovendi') work outside the universities and are financed by their employers or are unpaid for their dissertation work. The PhD students are supervised by at least one 'promoter' (a full professor). Usually, the internal PhD students are co-supervised by one or two 'co-promoters' (associate or assistant professors) who are responsible for the daily supervision of their research. Moreover, internal PhD students follow a formal training programme.

Dutch universities have organised this formal training in such a way that the general part (e.g., writing in English, giving presentations, general methodological courses) is taught by the universities, while the (domain-)specific parts are taught by the National Research Schools (NRS), especially ISED (Institute for the Study of Education and Development) and the Interuniversity Centre for Educational Sciences (ICO). According to the committee, these

NRS represent an important and highly attractive system for guaranteeing a high level of PhD training relatively independent of the local situation in the participating universities. Recently, the policy of most universities seems to have awarded more prominence to the PhD training offered within the university and, thus, to reduce the importance of these NRS. Particularly the formal tasks (e.g., approval of PhD project, keeping track of the progress of the PhD project, etc.) of the NRS have been decreased, while they have kept their offer of training activities (e.g. special courses at the PhD level, scientific meetings, international summer schools) as well as their network function.

The majority of PhD students who are currently active in the programmes under review completed a research master before they started their PhD training. It is not the task of this committee to assess the quality of these research master programmes, but the overall impression gained was that they generally provide an excellent preparation for a PhD project. However, it is questionable whether this warrants the planned reduction of the typical four-year PhD projects to three years.

The available tabular data about the time it takes PhD students to finalise their PhD thesis were generally difficult to interpret and evaluate, first, because not every PhD student had a four-year contract and, second, because many PhD students did not have a full-time position or did not work continuously due to illness, maternity leave, etc. Moreover, in some institutes part of the PhD students' available time (e.g., 10 or 20%) has to be devoted to other academic work (e.g, teaching assistance). This being said, it is clear that of the PhD students who started their training in the period 2003-2007, only a minority of 30% graduated after 4 years, while another 26% did so after 5 years and 10% after 6 years, implying that about one-third of those who started in that period either took even more time to graduate, did not finish at all or discontinued.

Completion status	Frequency	Percentage
Graduated after 4 years	53	30%
Graduated after 5 years	45	26%
Graduated after 6 years	18	10%
Graduated after 7 years	4	2%
Not yet completed	42	24%
Discontinued	14	8%

Table 2: Frequencies and percentages completed and discontinued dissertation projects of 176 internal PhD students who started their research between 2003 and 2007, computed across the six institutes.

There are also clear differences in completion times between institutes, although these differences may to some extent be due to differences in the way these numerical data are collected and presented, and/or to differences in the time that is needed to organise the actual defense once the PhD thesis is ready. These differences between institutes are also due to different practices and cultures with respect to timely PhD finalisation (e.g., giving rewards to those who finish in time versus giving an extra contract to those who need more time, etc.). In recent years, institutes have started to pay more attention to these completion rates. Nevertheless, given this numerical data, the committee shares the concerns expressed by

most institute managers and programme leaders about the reduction of the duration of a PhD project from four to three years (in line with with international developments).

The practice of submitting the PhD thesis in the form of a collection of articles (some of which have already been published, are under revision or submitted to an international journal) has become the rule in the fields of pedagogy and educational sciences. The committee has looked at a representative set of PhD theses and noticed that in many cases, they contained one or more papers that were published or in press in high-quality international journals or books. Clearly, these publications constitute a substantial part of the international publication output of the programmes, and, consequently, contribute significantly to the institutes' and programmes' productivity. Only in the subdomains of theoretical and historical pedagogy did some PhD theses still take the form of a monograph.

The interviewed PhD students' comments about the value of their training were generally very positive. They highly appreciated the courses, conferences, and contacts with PhD students from other universities provided by these NRS. Therefore, the committee makes a plea for upholding this well-established system of national schools, which play an important complementary role to the training provided within the universities. Clearly, these NRS also contribute substantially to research exchange and cooperation among senior staff, too. PhD students were also generally very satisfied with the part of the training provided within the institute, the opportunities for scientific and social contacts with their peers, their supervisors, and other staff members, the research facilities, and the various kinds of financial support that are available to them to conduct their research, attend conferences and visit research centres abroad. Finally, the PhD students that the committee spoke to see themselves as well-respected and well-supported as (junior) partners in a dynamic and multi-layered scientific network rather than as solitary students who work on a thesis under the supervision of a promoter – which is laudable.

According to the information provided by the institute and programme leaders, many PhD graduates found employment as postdoc researchers in the Netherlands or abroad, and as assistant professors at their own university, another Dutch university or a university abroad. However, none of the institutes appeared to systematically keep track of the further (academic) careers of their PhD students. Such information may be useful in assessing the institutes' 'long-term productivity' as well as providing important data for institutes on how they are meeting longer-term training goals.

In sum, from an international perspective, the Dutch approach to the training of PhD candidates, and particularly the internal PhD students, is of very good quality.

Overall evaluation

Based on the current assessment, the committee is generally very positive about the quality, productivity, social relevance and viability of the pedagogical and educational research in the Netherlands. In most cases it was also impressed by the amount and level of improvement over the assessment period. All programmes are nationally and even internationally visible and well recognised for at least certain parts of their work; and in some cases and/or for some domains or aspects, the research clearly reaches the level of international excellence.

Assessment at Institute and Programme level

3. LEIDEN UNIVERSITY

A. Institute level

1. The institute

The Institute of Education and Child Studies is one of the five research institutes of the Faculty of Social and Behavioural Sciences at Leiden University (LEI). It specialises in the empirical study of normal and deviant child upbringing and development.

Research at the Institute of Education and Child Studies is classified according to children's social contexts. It is subdivided in three research programmes:

- Child and Family (LEI1);
- Child and Educational Settings (LEI2);
- Child Welfare Services (LEI3).

According to its mission statement, and in line with Leiden University's overall research policy, the institute considers an interdisciplinary approach as fundamental to understanding issues of education and child studies. In the period under review, Leiden University has emphasised this need for interdisciplinary research by identifying eleven multidisciplinary research profile areas. The institute participates in two of them:

- *Brain function and dysfunction over the lifespan*, which brings together cognitive and biomedical neurosciences in the multidisciplinary setting of The Leiden Institute for Brain and Cognition (LIBC, set up in 2006) and the Leiden Center for Translational Neuroscience (LCTN, 2008)
- *Health, prevention and the human life cycle*, which focuses on issues of health and prevention and brings together research groups from the Faculty of Social and Behavioral Sciences and the LUMC.

Besides interdisciplinarity, other central elements of the institute's – very ambitiously defined – mission statement are: conducting innovative high-quality research, contributing to evidence-based practical interventions and recommendations, disseminating its research findings to the general public, and providing an excellent training of the next generation of scholars.

The institute is governed by a Board of professors (including the coordinators of the three research programmes), under the rotating chairmanship of a scientific director. During the current assessment period, responsibility for research has shifted from the Faculty to the Graduate Schools and the scientific directors of its research institutes. In the opinion of the committee, the organisational structure is balanced and well defined. The management is adequate.

2. Policy on scientific quality and academic reputation

To testify to the institute's scientific quality, the self-evaluation report mentions several general indicators of its academic reputation. The committee notes that the institute scores generally high to extremely high on these 'classical' signs of academic repute (i.e., keynote lectures at international conferences, participation in foreign PhD committees, editorships and memberships of editorial boards of important international and national journals, participation on boards of scientific and professional organisations, etc. It has also established

that institute members, particularly from LEI1, have won a high number of national and international prizes and awards, including some prestigious ones. More details are provided in the programme evaluations.

Evidently, this is an institute that has built a stellar programme of research and training. It is known and respected throughout the international research community for its leading edge work on developmental, educational and intervention research. Its flexible move to understanding the multi-level models including the genetics and epigenetics of child development is a sign of its growth and leadership in the field. Its emphasis on how such multi-level risk factor research can help to explain the effects of both social-emotional and reading intervention outcomes is cutting-edge and has international visibility.

The committee sees its positive assessment of the institute's quality and academic reputation reflected in the citation analysis by Prins, in which the institute scores very well, mostly with citation scores far beyond the reference standards for the distinct (sub)disciplines. The institute has clearly succeeded in supporting the ambitions of its researchers to publish in high-impact journals. For LEI1 these citation scores are exceptionally high. It should be pointed out, however, that it is this programme that strongly argued against the validity of this citation analysis based on Google Scholar (instead of Web of Science).

3. Resources and resource policy

Staff numbers at the Institute of Education and Child Studies have increased over the review period: from 84 persons (equal to 27.65 research fte) in 2006 to 92 persons (equal to 39.64 research fte) in 2011. This was largely due to an increase in non-tenured staff (mostly PhD candidates). The number of tenured staff members first went up (from 13.45 fte in 2006 to a peak number of 17.07 fte in 2008), but dropped in the last two years of the review period (to 11.54 fte in 2011). This is particularly due to the reduction of research fte for tenured staff in the LEI2 programme.

The increase in total research fte was paralleled by an increase in total funding, from €2,159,000 in 2006 to €4,192,000 in 2011. According to the committee, the remarkable increase in external funding is convincing evidence of the success of the preferred research policy of the institute. The ratio between direct funding and research grants has shifted over the review period. While the share of direct university funding in the institute's annual budget went into steep decline (from 62% in 2006 to 24% in 2011), the share of research grants and contract research went up (from 17% and 11% in 2006 to 51% and 16%, respectively, in 2011). There are notable differences between the three programmes in this distribution. In total, the three programmes received €7,893,750 in indirect funding (research grants and contract research) over the review period. This is equivalent to an earning capacity of €96,359 per tenured research fte per year.

Substantial grants were received from the European Research Council (ERC), the Ministry of Justice, the Ministry of Education, the Royal Netherlands Academy of Arts and Sciences (KNAW), the Netherlands Organisation for Scientific Research (NWO), national foundations, special purpose foundations and local community funds.

The institute's infrastructural resources are very good. The institute has made great efforts to keep abreast with the latest research technology and actively attempts to find additional funding to purchase and maintain essential equipment and to offer its researchers good access to this large and advanced set of facilities. In this respect, it is worth noting that the institute has been successful in acquiring a kind of 'beta status' during the assessment period, which

led to additional direct funds especially to improve the research facilities and infrastructure. However, this increasingly strong reliance on costly, new, technological infrastructure makes the institute somewhat vulnerable to budget cuts, as the funding market is getting ever more competitive and volatile.

4. Productivity strategy

Over the last six years, the research staff of the institute produced 727 academic publications (a yearly average of 3.7 per research fte), 478 of which were published in refereed international journals (with a more or less stable yearly average of 2.4 per research fte). There was a total of 212 professional publications, which means a rather low annual average of 1.1 per research fte. Both academic and professional productivity vary considerably between the three research programmes.

In total, over the review period, the programmes produced 37 internal PhD theses and 19 external PhD theses. This means that there was an average of 0.7 PhD theses per tenured fte, which is relatively low. Again, there are great differences in the number of PhD students between the three programmes, with one programme (LEI3) scoring particularly low.

5. Societal relevance

The committee has established that the institute's contribution to society is generally very high, and this holds for all three programmes. Even so, there does not seem to be a clear and strict policy in this respect at the institute and/or programme level. Accomplishments in this area are for the most part the result of the contributions of individual researchers or research groups. One notable element in terms of institute policy is that it has installed a special award for the popularisation of scientific knowledge.

Many current and new societal problems, such as immigration, child maltreatment, maternal smoking during pregnancy, and poor reading, which are at the heart of the institute's research agenda, call for scientific knowledge and evidence-based interventions and recommendations that the institute can provide.

The institute also contributes more directly to social issues, by developing numerous, evidence-based, practical recommendations, interventions, guidelines and policies. Dutch and even foreign policy makers, foundations, and professional organisations have asked members of the Institute to collect and present data about these and other social issues, and the committee has seen clear evidence of actual utilisation by these external target groups.

Part of the institute's mission is to disseminate research findings to the general public. The institute's members frequently give popular lectures, present results to specific stakeholders (e.g., adoptive parents, teachers, parents of disabled children), give interviews, and write articles and books for a professional and popular audience, etc.

Finally, the institute also received various signs of demonstrable recognition from external target groups, such as public awards and prizes, invited memberships of councils of national and international educational, social and health organisations. The fact that the institute obtains a substantial share of its funding from contract research, and that eight of its researchers have held or still hold an endowed professorship, can be seen as additional signs that its societal relevance is acknowledged and appreciated outside of academia.

6. Strategy for the future

Given the very good to excellent quality, productivity and societal relevance of its research, its successful focus on interdisciplinary research that is well-embedded in the university-wide profile areas, and its continued emphasis on fundamental research with an eye to practical applications, the Institute seems very well placed to make a highly significant contribution to tomorrow's (inter)national science and society. Moreover, the issues on which the institute focuses will not disappear overnight, and funding agencies have increasingly acknowledged the institute's expertise on these issues.

Even so, the self-evaluation report identifies some threats to the institute's viability. First, the financial reliance on competitive and volatile external funding reportedly makes the institute vulnerable to budget cuts, especially since it uses costly new technology. Second, given the exceptionally strong reputation of some of its senior members, the institute faces a challenge to find worthy replacements. It has become clear to the committee, however, that this threat has effectively been mitigated during the review period. The institute has made plans to begin to anticipate upcoming retirements. In 2012, it hired nine new assistant professors, who can be expected to replace the present senior researchers in the foreseeable future. Several of these younger scholars already play an important role in the governance of the institute and its programmes. This forward planning is notable and provides continuity for the quality and productivity of the institute.

The strategy for the future, as set forth in the self-evaluation report, contains three main elements. The first is to intensify collaboration with international researchers from various disciplines to address fundamental and applied issues. Clearly, this strategic point applies more to some programmes than to others. Second, to focus more on acquiring international (especially European) funding, for which a high-level European network is a prerequisite. Arguably, achieving these two strategic goals may be helped through the institute's active participation in the university's research profiles. The third and final element is to make sure that the expected reduction of the PhD time from 4 to 3 years does not lead to a loss of quality, by increasing the selection pool of research master students and setting high standards for PhD candidates. How successful this strategy has been so far is not entirely clear to the committee. During the site visit, the insufficiently productive recruitment of young local post-graduate students was acknowledged as a problem by at least one programme.

While the committee largely agrees with the course for the future as set out by the institute, it would like to propose one additional element. In general, the committee has found that, while LEI1 and LEI2 show some integration across the institute, there is the possibility of better integration, especially with LEI3. While it is exemplary that the institute as a whole has had a strong, long-term focus on issues related to culture, immigration, and youth development, the committee considers this an important part of Leiden's work and of its high social and policy relevance. Thus, it is somewhat surprising that, in spite of some inter-cultural work in each of the three programmes, there appear to be few or no relations between these programmes at either the faculty level or in the PhD training programme. This would appear to be a lost opportunity for cross-programme integration. The committee therefore recommends improving internal coherence, not just between programmes, but also within LEI3, which – at the moment – contains two unrelated lines of research.

7. PhD training and supervision

Students can follow a two-year research master's degree programme, Developmental Psychopathology in Education and Child Studies, to prepare for a PhD career at the institute.

Internal PhD candidates are currently employed on the basis of a four-year 1.0 fte contract, but there is pressure to switch to three-year contracts. As stated above, the institute leadership fears that this reduction of the PhD time may lead to a loss of quality, a fear that is shared by the committee: given the institute's focus on multidisciplinary training, dissertations will likely continue to take four years (at least). The committee was happy to hear that the institute is taking measures to prevent the loss of quality.

The PhD training is adequately organised. In addition to training at Leiden University via the local graduate school, PhD candidates participate in national research schools (NRS), mostly ISED, the Institute for the Study of Education and Human Development. Through ISED, the Institute's PhD candidates can take courses at various Dutch universities.

Courses in scientific English and the presentation of research findings form part of their PhD training at the local graduate school. The major aim of this graduate school is to integrate students into the Leiden scientific community and to acquaint them with its rules, ethical standards, and regulations.

PhD candidates work on the basis of an approved research plan written together with two or more senior researchers, who act as supervisors.

PhD students are required to present their results in English at weekly meetings at the programme level, at faculty meetings, and at international conferences. A three-yearly prize is awarded for the best PhD thesis.

Out of the 35 funded PhD students enrolled between 2003 and 2007, 23% had graduated after four years. After five years, another 14% had graduated, and after six years another 11%. After six years, only 48% of PhD students had finished their projects, while 35% was still working on them, and 17% had dropped out. Comparatively speaking, these are rather poor graduation rates. According to the self-evaluation report, the relatively long completion time can be explained by the fact that several of the candidates postponed completion for personal reasons (e.g., pregnancy, illness, 0.8 appointment for 5 years) and by the fact that there is a considerable delay between approval of the dissertation and the public defense. In the meantime, the majority of the unfinished PhD theses have been successfully defended.

From its interview with a selection of PhD students, the committee got the impression that they are generally very pleased with the content and the structure of their PhD training, and with the quality of their direct supervision.

Many PhD graduates from the past six years found employment as postdoc researchers in the Netherlands or abroad, and as assistant professors at Leiden and other Dutch universities.

B. Programme level

The committee comes to the following overall programme scores for quality (Q), productivity (P), relevance (R), and vitality/feasibility (V) for the three programmes of The Institute of Education and Child Studies of Leiden University:

Code	Programme name	Q	P	R	V
LEI1	Child and Family	5	5	5	5
LEI2	Child and Educational Settings	4	4	4	4
LEI3	Child Welfare Services	4	4	4	3.5

The detailed assessment for each programme follows in the next section of this report.

Programme LEI1: Child and Family

Programme director: Prof. Marian Bakermans-Kranenburg, Prof. Femmie Juffer

Research staff 2011: 4.3 fte tenured, 16.7 total fte

Assessments:	Quality:	5
	Productivity:	5
	Relevance:	5
	Viability:	5

Brief description

This programme examines the role of parents and other caregivers in children's development. Factors influencing development are considered very broadly, including genetic and hormonal influences, family relationships, and cognitive functioning.

Quality

By all measures, the quality of the programme is excellent. Its three directors are internationally renowned scholars, who collaborate with other leading scholars from around the world. The research is well-integrated and coherent. Much of it is very frequently cited and published in excellent international outlets. Generally, it is regarded as being among the best in the world. Evidence of the programme's outstanding reputation is also found in the many honours and academic prizes received by its faculty. Before, during, and following the review period, the programme has engaged in cutting-edge research.

Some of the primary studies conducted by the programme are longitudinal, and thus require some time to be published. The strategies for publication ensure a high level of productivity nonetheless, by also encouraging papers with shorter turn-around, such as papers studying G x E (Gene X Environment) interactions and meta-analyses. In terms of bibliographic outcomes, the programme scores at the excellent level, including many highly cited publications.

The programme was very successful in attracting external research funds, with funding from the ERC, the Ministry of Justice, KNAW, and NWO. Funding increased significantly during the review period (from €388,000 in 2006 to €1,443,000 in 2011).

Productivity

Considering all measures of productivity along with the productivity strategy, this programme is judged as excellent. The total number of academic publications produced during the review period was 275 (which amounts to an annual average of 3.29 per research fte). The number of refereed articles in English journals is 201 (2.4 per research fte). The number of professional publications is reasonable (115, 1.38 per research fte). A remarkably large number of papers are published in high-impact journals. Even within these high-impact journals, the papers are cited very highly. This is also demonstrated by the exceptionally high *h*-indices of the three programme leaders, and indicates an excellent productivity strategy. Of the 26 theses defended during the review period, 22 were internal and 4 were external projects. This amounts to 0.91 theses per fte for tenured staff.

Societal relevance

The programme has very strong relevance to society nationally and internationally. The body of scholarly work has an extraordinary impact nationally and internationally because it focuses on the needs of high-risk children and families and attracts attention well beyond academia.

Staff work to provide systematic, empirically based answers to topics of high societal relevance (e.g., concerning abuse, adoption). Examples include the Netherlands' Prevalence Study of Maltreatment of Youth, which was cited extensively in parliamentary documents, newspapers, and magazines. In addition, the programme has been involved in developing an evidence-based treatment for high-risk children that is disseminated nationally and internationally.

Viability

The programme's viability is considered excellent because of its internationally leading role, its involvement in cutting-edge research ideas, its flexibility in adjusting as new ideas and technology emerge, and its extraordinary leadership. The programme has excellent support at the university and institute level. The university provides funding for infrastructure development, and the hiring of new permanent staff members is already planned, which provides concrete evidence of university support. The tenure of Prof. Van IJzendoorn, who is several years from the usual retirement age, has been extended, which is critical to the programme's success. Prof. Bakermans-Kranenburg is nearly two decades from the usual retirement age, which will ensure continuity. The programme's impact has increased during the study period, and given the strategies in place, there are good reasons to expect that it will continue to do so. Also, by hiring new staff, the continuation of the current successful research lines will be guaranteed.

Conclusion

The programme is judged to be excellent with respect to the quality of the research conducted, the quantity of output, the relevance to society, and its viability. The programme directors are international leaders in the field who make highly significant contributions to a number of areas of research.

Programme LEI2: Child and Educational Settings

Programme director: Prof. Adriana Bus and Prof. Paul van den Broek (since 2007)

Research staff 2011: 3.4 fte tenured, 10.1 total fte

Assessments:	Quality:	4
	Productivity:	4
	Relevance:	4
	Viability:	4

Brief description

The programme focuses on the specific mechanisms by which risk and proactive factors exert their effect on learning. The main expertise concerns learning and instruction of literacy and, to a lesser extent, other academic skills. During the assessment period, the programme has undergone some reduction in fte's as well as a substantial change in focus. The present programme consists of three interconnected strands: literacy and reading comprehension, learning-related interventions, and the role of teachers in fostering learning and skill development.

Quality

The ambition with regard to quality is very high, as shown by publications in forums such as *Science* and *Psychological Bulletin* by the two senior researchers. The intense search for the most important factors associated with literacy or reading comprehension has led to the use of methodologies such as meta-analyses and tools to quantify comprehensibility. Both subprogrammes are very ambitiously directed, which means that they have generated a lot of citations and admirable comments from research communities working in related areas. The key journal articles of the senior professors are published exclusively in very high-impact journals. A substantial portion of the publications have emerged as invited articles in edited books.

Senior members have been very active in editorial duties and on international evaluation and advisory committees, and have been repeatedly rewarded for their work, e.g. in the form of prestigious membership of the KNAW.

The programme was successful in pursuing external sources of funding, not only at the national level but also internationally. However, the absolute amount of indirect funds, and the amount of indirect funds acquired per senior staff fte are rather small; accordingly, the proportion of these indirect funds in relation to the programme's total budget is quite low compared to other programmes in this assessment, although the self-evaluation report clearly shows a steep increase in the number of national and international grants obtained during the last two years of the assessment period.

The general picture concerning quality is compromised by changes in staff. The newest and youngest members have not yet had time to demonstrate achievements, which could complete the picture of the whole programme in an adequate way. Also, new members are opening up research areas such as brain research associated with child development and learning and curriculum-based measurement, which have no strong prior history in this research environment. These changes in staff lower the opportunity to show fast results, and alter the internal coherence of the programme.

The programme actively pursues international collaborations. Theoretical models and methodological tools developed in the domain of reading comprehension are frequently used by other researchers worldwide (e.g. the Landscape model of reading comprehension as well as a simulation tool that facilitates quantification of the qualities of written material based on this model). Similarly, literacy research, which includes collaboration with researchers in developing countries, is an interesting aspect of the international work.

Productivity

The publishing activity is substantial, with a total of 211 academic publications (an annual average of 3.18 per research fte), 124 of which appeared in international journals (1.87 per research fte). A larger than typical proportion of academic publications appears in other outlets such as book chapters and proceedings. The number of professional publications (a total of 51) is rather small (0.77 per research fte). During the evaluation period, 11 internal and 12 external PhDs defended their theses, resulting in a total of 0.78 PhD per tenured fte, which is rather moderate.

Societal relevance

The programme has resulted in various products, including reports for professional organisations, publications in professional journals and series, and publications for the general public. Many key research topics seem to be oriented to outcomes that lead to, or support, applications which can have a widespread use. Most of these assist developments associated with literacy acquisition in general, independent of specific language environments. In this respect, the programme has produced applications that are used in many Dutch schools (e.g., software for literacy and reading) and elsewhere (e.g., ‘curriculum-based measurement approach’, which is used internationally). Programme members have been actively participating in guiding teachers and parents in the many ways to support learning of academic skills. Moreover, they have been very active in contacts with educational stakeholders and participating on advisory boards of educational organisations, the media, etc. A final activity with regard to societal relevance is the work conducted to support young researchers in developing countries to help them to acquire expertise that they can use to enhance reading instruction in their countries.

Viability, feasibility and vision for the future

The preference for research questions relating to the real-life context of children’s learning of basic skills ensures that the impact of the programme is strong, while making it well-suited to attract financial support. The most visible research focuses on a relatively narrow part of children’s learning and educational settings. This is understandable and wise due to the small size of the programme in terms of manpower. The way in which the topical areas have been chosen, and their success in attracting international attention, inspire confidence in the vitality and viability of the programme. Two manifest threats, which the programme leaders have already started to address successfully during the last two years of the assessment period, are (1) the difficulty in attracting good PhD students and (2) the high dependence on direct funds. On the other hand, programme leaders have been rather summary in announcing their visions about future plans. It remains unclear to the committee how they envisage continuity to bridge their retirement.

Conclusion

The programme has experienced major changes in personnel, which affect the whole picture. The senior staff has been focussed on important research, and the results have been very good. The prospects are also very good in the specific areas in which they have had a great impact on international research communities: literacy and reading comprehension. The

excellent balance between scientific and societal goals is another strength. Renowned educational researchers who were recruited from the US with a view to add to the scientific strength and viability of the programme have not yet had time to make their mark on the programme. How well the research strands will be integrated is still a slightly open question, although the two most central research foci (literacy acquisition, reading comprehension and interventions) are closely connected conceptually.

Programme LEI3: Child Welfare Services

Programme director: Prof. Hanna Swaab and Prof. Paul Vedder

Research staff 2011: 3.9 fte tenured, 9.8 total fte

Assessments:	Quality:	4
	Productivity:	4
	Relevance:	4
	Viability:	3.5

The stated mission of this programme is to study children and adolescents whose behavioural, emotional, and social functioning are seriously challenged. The objective is to improve their health by utilising the family, the school, or more specialised facilities. Studies within this programme focus on the brain-behaviour relationship and the early identification and manipulation of those mechanisms that modulate developmental risk.

Typically, the projects in this programme address the complexity of the interaction between brain (dys)function, (neuro)cognitive function, behavioural symptoms, and environmental factors. The aim is to improve health services for severely or mildly disturbed children or children and adolescents at risk. The programme employs longitudinal studies to unravel developmental pathways to psychopathology and to appraise the adequacy of early interventions in children. Examples of such projects include studies of children with autism spectrum disorders (ASD), genetic disorders or ODD/CD, and pre-psychotic adolescents, and studies into the effects of environmental influences on child development (maternal smoking during pregnancy, the effects of perinatal asphyxia, cancer treatment, and sleep deprivation, and the effect of cultural influences and migration).

Quality

The programme has substantially changed in the past five years into a programme primarily focused on ‘mechanisms of social dysfunction’. The great majority of the research is on brain-behaviour relations in DSM-type conditions of children and youth (autism, psychosis, ADHD, aggression, etc.). The primary focus of one sub-component of the programme is on ‘developmental psychopathology’. This sub-programme has grown quickly but lacks a clear conceptual model that ties together its component projects.

A second, continuing strand of the programme concerns cultural issues regarding immigration, schooling, and youth development. This is an important line of research, which has also led to a strong rate of publication in leading journals in the field and a very viable research direction.

This programme of research has shown substantial development over the last five years. Its quality is recognised by its high output of international journal articles in well-recognised journals. Further, the programme has been very successful at obtaining research grants and contract research assignments in both sub-areas of interest (with an impressive increase in success with respect to the former category).

A substantial concern is the lack of coherence of the programme. This is reflected in the lack of interaction between PhD students across the programme. Further, the term ‘Child Welfare Services’ does not appear to reflect the programme’s areas of focus.

Productivity

The programme has been very productive. There has been an upsurge in academic research output: 247 academic publications (5.01 per research fte), 152 of which are refereed articles in English (3.10 per research fte). The number of professional publications is reasonable (46, i.e., 1.36 per research fte).

The number of PhD theses is very low with only 7 completions during this period (4 internal and 3 external; 0.29 per ft for tenured staff), and this is of substantial concern. Thus, the major concern regarding productivity lies in the PhD programme.

Societal relevance

Both aspects of the programme score high on measures of social relevance. The sub-programme on social dysfunction has strong connections to clinical settings and schools, and the findings are of substantial public interest. There is strong outreach to parents through services provided in clinical settings for families with children with behavioural disorders and other disabilities. The sub-programme on immigration, cultures and children's services provides strong outreach to communities around the country. It also offers on-going consultation with government agencies and has had substantial effects on policy discussion and decisions regarding youth.

Viability, feasibility and vision for the future

The programme is relatively new, but the results thus far are very promising. Research in both sub-programmes areas is vital. The publication output and grant acquisition are high and likely to show continued growth. The research areas are of substantive interest to Dutch and international agencies and should be competitive. However, the committee has significant concerns regarding the viability of the programme itself, given that it is splintered and lacks coherence.

Conclusion

This is a newly designed programme. Although its faculty have done very well with regard to quality, productivity, and societal relevance, the programme lacks coherence and vision. Still, both the bio-psycho-oriented research on developmental psychopathology and the more applied youth research on immigration, acculturation and youth development are viable and productive areas of great interest. Each area of research is strong and will likely continue to make important contributions to both science and practice in child development.

4. UNIVERSITY OF GRONINGEN

A. Institute level

1. The institute

The Nieuwenhuis Institute for Educational Research (NI) is part of the Faculty of Behavioural and Social Sciences of the University of Groningen (RUG). It consists of four research programmes, three of which are featured in this research review. These three programmes essentially coincide with the three programmes that were assessed in the previous review, although some members have since moved from one programme to another. These programmes are:

- Education in Culture (RUG1), which focuses on the educational process as a whole and in its diversity, with history and philosophy of education and sociology of youth as the three main subdisciplinary perspectives;
- Research and Evaluation of Educational Effectiveness (RUG2), with expertise in school-effectiveness research, evidence-based education and longitudinal research in education;
- Developmental and Behavioural Disorders in Education and Care: Assessment and Intervention (RUG3), which specialises in care for children and youngsters with learning and educational problems, disabilities, and emotional and behavioural problems, with a focus on severe and multiple handicaps or problems.

The overarching objective of the Nieuwenhuis Institute is to increase the knowledge base on education and to contribute to the solution of problems within educational practice. This objective is established through a combination of theory-related and practice- and policy-oriented research.

Although every programme within the NI shares this twofold objective and strategy, there are differences in emphasis, as will become clear in the assessments at the programme level.

The institute operates quite independently from the faculty, although formal responsibility lies with the Faculty Board. Both the Graduate School and the Research Master Behavioural and Social Sciences are organised at the level of the faculty. The institute's basic management philosophy is that of collegial management. Long-term strategic decisions are taken unanimously. The institute's director, who is appointed by the faculty board for a period of three years, takes operational and mid-term strategic decisions after consulting the advisory board. All programme leaders are *ex officio* members of this advisory board. Accountability for the orientation and internal functioning of the research programmes rests essentially with the programme leaders.

During the previous assessment period, the Nieuwenhuis Institute was still recovering from a drastic reorganisation, involving the termination of one programme and a serious staff reduction in 2005, due to financial problems at the department level. At that time a rigorous strategic plan was made to make the department financially healthy again, which has been accomplished in the meantime.

2. Policy on scientific quality and academic reputation

In line with the RUG tradition, the institute leadership intervenes as little as possible in the orientation and functioning of its various research programmes. This policy is based on a long-standing and strong belief in the academic freedom and autonomy of professors, in the

viability of rather small, focused and stable programmes, and in the scientific, societal and educational adequacy of the chosen research topics and agendas. The result of this policy is a set of rather small, stable and rather independent programmes, which are strongly embedded in the traditional (sub)discipline(s) of educational sciences (such as history and theory of education, ‘orthopedagogy’). Each of these programmes has its own unique and nationally – and to some extent internationally – well-known ‘selling points’.

At the same time, the self-evaluation report shows that the institute and programme leaders are well aware that this focus on autonomy, internal (sub)disciplinary coherence, and stability may involve a risk, especially since some programmes are situated in rather small and specific niches of research – nationally but also internationally – with somewhat atypical research and publication cultures and practices. This holds particularly for RUG1, but to some extent also for RUG3.

While respecting the institute’s structure and culture, a number of strategic initiatives and actions have been taken over the past few years. They were intended to increase the scientific quality and reputation, at both the institute level and the programme level, after the previous research assessment of 2007 (see below). All new academic personnel are now hired on a six-year ‘tenure track’ contract, which includes clear and rather demanding criteria for success, in terms of publications, grant acquisition, organisational and supervision skills, and teaching. The already appointed personnel is evaluated according to these same criteria.

Generally speaking, the committee feels that the scientific quality and academic reputation of the Nieuwenhuis Institute is good to very good, even though few papers appear in top tier international journals. The self-evaluation report lists various pieces of evidence that the research is nationally and internationally visible and makes valuable contributions to the international field. This is reflected in invitations to present keynote speeches at international conferences, in active participation in international societies and organisations, conferences and projects, in visiting professorships, in editorships of international academic journals, in an increasing number of scientific grants obtained, in participation in international research projects, and also in a few prizes and awards. For more detailed descriptions and evaluations, the committee refers to the programme assessments below.

To testify to the quality of the research, the self-evaluation report points to the outcomes of the citation analysis by Prins (2012), which showed that the better part of the institute’s scientific output is at or above the international citation level. Compared to the other institutes, the results of this analysis are rather modest, particularly for the RUG1 programme. However, as argued in the self-evaluation report, this may be mainly caused by the fact that the RUG1 research programme is relatively small and situated in a smaller scientific niche, with an atypical publication culture.

3. Resources and resource policy

After a considerable decline in personnel in the aftermath of the 2005 reorganisation (see above), research staff numbers at the Nieuwenhuis Institute have significantly increased: from 54 (equal to 19.4 research fte) in 2006 to 83 (equal to 36.4 research fte) in 2011, largely due to an increase in non-tenured staff in RUG2 and RUG3. According to the self-evaluation report, this increase is a consequence of the start of a new bachelor’s programme (for primary education teacher training), the availability of new direct funds for PhD and postdoc positions as of 2009, and an increase in research grants and contract research. Generally speaking, the number of tenured staff members was stable during the review period (although

the associated fte's rose slightly, from 9.6 fte in 2006 to 11.7 fte in 2011). At RUG3, the number of tenured fte's gradually increased by about 1.25 fte.

Over the review period, the total annual budget of the Nieuwenhuis Institute has more than doubled, from €2,061,000 in 2006 to €4,234,000 in 2011. RUG1 is by far the smallest programme in terms of fte, and received between 8.2% and 15.5% of the funding, while RUG3 obtained between 20.8% and 43.5%. Remarkably, RUG2 received between 48.4% and 68.4%, even though this latter programme is not substantially larger (in terms of research fte for tenured staff) than RUG1 and considerably smaller than RUG3.

The ratio between direct funding, research grants and contract research has fluctuated slightly over the review period, but overall the categories are of similar size. In 2011, the institute's annual budget was made up of 30.2% direct funding, 39.1% research grants and 30.6% contract research (in 2006 this was 36.9%, 34.1% and 29.1%, respectively). So, comparatively speaking, there is a strong emphasis on contract research, both in absolute and relative terms. In total, the three programmes received the very respectable amount of €10,853,487 in indirect funding (research grants and contract research) over the review period. This very good result seems particularly due to the earning capacity of RUG2. This makes for an overall earning capacity of €166,567 per tenured research fte per year. In the opinion of the committee, the clear upward trend in the acquisition of external money is a positive element, given that in the previous review the rather small grant acquisition capacity was considered a major weakness of most programmes of the institute.

4. Productivity strategy

Over the last six years, the research staff members of the three programmes produced 604 academic publications (a yearly average of 3.6 per research fte), 268 of which were published in international (English as well as non-English) refereed journals (a yearly average of 1.6 per research fte). Furthermore, the institute produced a total of 575 professional publications, amounting to a very high annual average of 3.4 per research fte.

Like grant acquisition, academic productivity (especially in terms of English and other international journal articles) has considerably improved compared to the previous assessment period. The tenure-track system that is in place has encouraged faculty to increase not only its academic quality and grant acquisition capacity, but also its academic productivity, while the productivity of professional publications, which was traditionally one of the institute's strengths, has remained strong.

The small number of PhD theses was another matter of concern in the institute's previous assessment. There has been some improvement with respect to that point, and the upward trend certainly has not stopped: when comparing the first three years of the review period (2006-2008) to the last three years (2009-2011), the number of completed PhD theses has doubled. In total, 40 PhD theses were completed over the review period, which amounts to an average of 0.6 completions per tenured research fte. This is, however, still rather small in comparison with most other institutes.

5. Societal relevance

The potential societal relevance of the institute's research is very high, given the nature of the themes that are at the forefront of the programmes' research agendas (diversity in education, evidence-based education, education and care of children and youngsters with severe and/or multiple problems). The committee established that this potential relevance is adequately effectuated in all three RUG programmes.

With respect to societal quality, the three research groups interact intensively and productively with various groups of stakeholders in society, and contribute directly to important issues and debates in society, ranging from sexual harassment and abuse in schools and residential care (RUG1), to evidence-based education (RUG2) and child-oriented social welfare (RUG3).

As far as the societal impact of the work is concerned, the self-assessment report contains various examples of how research has affected specific stakeholders and/or specific procedures in society, e.g., by developing concrete materials, modules, protocols about the three above-mentioned topics that are effectively and intensively used in educational and care settings in the country.

The institute scores generally high with respect to outreach. This is exemplified by the very high number of professional publications and various other ways of making its research results available to the general public (via articles and commentaries in newspapers, appearances and programmes for radio and television, weblogs, etc.).

The fact that the institute is able to acquire so many contract research assignments can be considered a clear demonstration of its recognition by its various external target groups. The institute has established strong and long-standing collaborations with many of them. One of the interesting benefits of the frequent and intensive collaborations with service providers is the possibility of recruiting external PhD students (particularly in RUG3).

6. Strategy for the future

Compared to the previous assessment, the institute has clearly improved in terms of quality and productivity of its scientific output, its external grant acquisition capacity, and the number of PhD theses, while keeping the traditionally high societal relevance and impact of its programmes. This is to a large extent due to the seriousness with which it has worked on the weaknesses and recommendations mentioned in the previous assessment.

This especially holds true with respect to personnel policies. In light of the continuous difficulties experienced in attracting young and talented staff from outside of the region to replace ageing and retiring staff members, the institute has introduced an ambitious and attractive tenure track system. To further boost academic quality and productivity, rigorous evaluation procedures for existing staff have been implemented. To guarantee a steady inflow of young researchers, it was decided to directly fund at least three PhD positions and one postdoc position per year.

In spite of these positive changes, the institute – generally speaking – still appears to lack coherence and a comprehensive vision. Its general strategy is to capitalise on the coherence, continuity and autonomy of its distinct programmes, which are relatively small in terms of staff size and in some cases suffer somewhat from the specificity of their scientific niches, and their accompanying research, publication, and citation cultures. According to the committee, this strategy involves some genuine risks. From interviews held with the institute management and representatives of the programmes, it concludes that staff members work mostly autonomously, and there is little integration across the programmes of the institute. The structure is decentralised, and decisions appear to be made primarily at the programme level, not the institute level. The lack of interdisciplinary collaboration is also reflected in the fact that cross-programme discussions among PhD students have not been instigated.

The committee further concludes that, unlike most other institutes, the NI makes little effort to acquire a prominent place in the RUG's interdisciplinary research priority areas, which

indeed may conflict with the value attached to the programmes' coherence, continuity and autonomy. The committee questions whether this 'conservative' strategy should be continued in the increasingly competitive and rapidly changing academic world.

The traditional *modus operandi* of the institute is also apparent from the fact that institute-level decisions appear to be made on the basis of teaching and teaching needs. Strategic efforts are largely directed at increasing student numbers, which – in time – will lead to an increase of the direct funds available for research. To the committee, a striking example of this teaching-based strategy is that both professorships that will become vacant due to retirement will be filled in the same teaching areas in order to retain students in the programmes. This decision reflects concerns regarding the institute's strategic planning and its potential to adapt to the rapid changes occurring in the social, educational and biological sciences.

In conclusion, the committee notes that, although it absolutely does not question the viability of the Nieuwenhuis Institute as such, further improvement towards excellence seems necessary, guided by the newly installed rigorous personnel recruitment and assessment procedures and by more systematic attempts to explore and compete on the market of European research funds. The institute mentions the external pressure to increase the programmes' scale as a threat to their future. Rather than pleading for enlargement of the programmes, the committee advises forming more cooperative relationships and strategic alliances, both within the institute and the university at large. This would genuinely enhance the institute's chances of obtaining important external research grants at the European level. Finally, it should be noted that the strategy to reserve direct funds for PhD and postdoc positions available for the research programmes deviates from the policy of most other universities and institutes. It may be an appropriate temporary measure to resolve certain short-term needs but, in the long run, one might consider if it is not a better strategy to stimulate the staff and the new generation of scientists to compete for these positions in the second and third money streams.

7. PhD training and supervision

At the faculty level, the Graduate School plays an important role in the training of PhD candidates. All PhD candidates are a member of the Graduate School and can take courses there. Additionally, PhD candidates participate in the National Research Schools (NRS) that provide specialised PhD training. For RUG1 and RUG3 this is ISED, the Institute for the Study of Education and Human Development, a joint enterprise of five Dutch universities. For RUG2, PhD training takes place within ICO, the Interuniversity Center for Educational Sciences.

Out of the 13 funded PhD candidates that enrolled between 2003 and 2007, only 8% graduated after four years. After five years, another 46% had graduated and after six years, another 8%. Thus, after six years, 62% of the candidates had finished their projects, while 38% had not yet completed them. Remarkably, there were no drop-outs. The institute claims that this completion rate and speed are generally satisfactory, but acknowledges that it can and should be improved. Having one's PhD students graduate within four to five years has been included as one of the tenure-track evaluation criteria of the staff.

The PhD students that the committee spoke to were generally pleased with their training and supervision. These students evaluated the national and local PhD training components as highly valuable and complementary. They enjoy the academic climate at the institute, the various possibilities to meet (in)formally with the other PhD students of their programme, and, to a lesser extent, with other RUG PhD students during courses at their Graduate

School (although they would appreciate courses in time management and writing grant proposals). The PhD students generally appreciate the availability of their supervisors and the quality of their supervision. They assessed the pressure to finalise their PhD thesis in time and to publish about their PhD research as appropriate.

B. Programme level

The committee comes to the following overall programme scores for quality (Q), productivity (P), relevance (R), and vitality/feasibility (V) for the three programmes of the Nieuwenhuis Institute for Educational Research of the University of Groningen:

Code	Programme name	Q	P	R	V
RUG1	Education in Culture	3.5	3.5	4	4
RUG2	Research and Evaluation of Educational Effectiveness	4	4	4	4
RUG3	Developmental and Behavioural Disorders in Education and Care: Assessment and Intervention	4	3.5	4	3.5

The detailed assessment for each programme follows in the next section of this report.

Programme RUG1: Education in Culture

Programme director: Prof. Jeroen Dekker and Prof. Greetje Timmerman

Research staff 2011: 2.3 fte tenured, 4.6 total fte

Assessments:	Quality:	3.5
	Productivity:	3.5
	Relevance:	4
	Viability:	4

Brief description

This programme focuses on the educational process as a whole, with history of education, theory of education and the sociology of youth as its main research perspectives.

The multidisciplinary research of the programme covers three research areas:

- Educational and cultural transmission through parenting and schooling;
- Children and youth at risk;
- Child and youth cultures from a generational and historical perspective.

Since 2006, the total research staff has increased in fte terms, mainly due to the recruitment of more non-tenured staff.

Quality

Following the recommendations of the last assessment, the number of research areas was reduced to those three listed above, with key appointments being made relevant to each. The programme has maintained strength and focus, and during the assessment period, important contributions have been made to research in several areas. However, the evidence suggests that there has not been an improvement in quality during this assessment period. The programme leaders seem to build on the reputation they have acquired, without providing strong evidence of a dynamic strategy aimed at raising the international position and impact of the programme and pushing it toward scientific excellence.

Nevertheless, very useful historical analyses have been made into educational policy and the circumstances and treatment of children at risk, valuable insights have been offered into the nature of Islamic education, and critical examinations have been made of youth culture and how it is handled in educational and social policy. The quality of this research is reflected by the quality and diversity of publications, which include monographs and edited collections as well as articles in international refereed journals.

Senior staff members have been awarded academic prizes for their work, have served on the editorial boards of international journals in relevant fields, have given invited keynotes in international conferences and have participated in international research projects. The international reputation of the research is good, and the programme generally makes a valuable and visible contribution in the international field.

Members of the programme have been successful in acquiring research grants from competitive and prestigious sources such as NWO, with the useful and applied value of their research being recognised by the awarding of several grants from government ministries. Overall, however, the grant acquisition capacity is small.

Productivity

In terms of bibliometric outcomes (including the Prins' analysis), the scores are not very good, which may to some extent be due to the content areas of the research not matching the content specification for some of the most prestigious educational research journals, as it lies in a niche between the humanities and social sciences. There is a relatively strong focus on national as compared to international contributions (e.g., the number of national compared to international books and book chapters). The programme is encouraged to use appropriate strategies to raise the level of its impact on the field and visibility in the wider international research community.

Given the small number of tenured staff over the assessment period as a whole, the research output has been good: 115 academic publications (5.3 per research fte), though only 37 of them are in international refereed journals (1.79 per research fte) and only 23 are international book chapters. This means a large proportion of the academic publications is in Dutch. The programme also produced a number of monographs, mainly in Dutch. On the other hand, the number of professional publications is very high (85, i.e., 4.12 per research fte).

The number of PhD students has remained fairly static throughout the assessment period and is still quite small. The number of PhD theses completed in the assessment period was 9 (6 internal and 7 external; 0.66 per fte for tenured staff). The completion rate was problematically low in the first half of the assessment period, but there has been a positive change during the second half. This suggests that the programme leaders and staff have taken the concerns and recommendations from the previous assessment reports seriously.

Societal relevance

The research clearly makes an impact on educational and social policymakers and practitioners nationally. Programme members have strong contacts with government ministries and a variety of influential bodies, including broadcasters and school boards, which are concerned with the education of ethnic minorities and children at risk. Thus, the findings of the research are being used to design training schemes in higher education and school counselling. There is regular contact between programme researchers, educational practitioners and the general public. Staff often makes active contributions to public debates.

Viability, feasibility and vision for the future

The programme's viability is good because of the distinctive contribution it makes to Dutch educational research as a whole and the recognised role of its members in shaping national educational policy and professional training, and because of its international visibility in history/theory of education research. The development of the programme supports the view that the criticisms and recommendations of the previous research assessment have been taken into account. Strenuous efforts are being made, quite rightly, to gain more external research funding and internationally visible output. On the other hand, the programme is confronted with some threats, such as the difficulty in recruiting talented young staff.

Conclusion

Overall, this is a good to very good programme, which is recognised nationally and internationally for making a distinctive contribution to educational research. It would be to the great detriment of the educational research community, internationally and nationally, if the original and distinctive lines of enquiry pursued within it were not maintained and developed. The recruitment of more young, talented staff and increased success in obtaining external grants and more international refereed publications would help secure its future.

Programme RUG2: Research and Evaluation of Educational Effectiveness

Programme director: Prof. Greetje van der Werf and Prof. Roel Bosker

Research staff 2011: 3.3 fte tenured, 13.6 total fte

Assessments:	Quality:	4
	Productivity:	4
	Relevance:	4
	Viability:	4

Brief description

This programme focuses on ‘what works why in education’. It has two key objectives: first, to explain why some students, teachers and schools do better than others, in terms of cognitive and non-cognitive outcomes; second, to integrate basic and applied research in such a way that its results are relevant and applicable for educational practice and contribute to more evidence-based education.

Since the previous review, the research topics have been reorganised into three subthemes:

- ‘Determinants of school success’ studies the effects of intrapersonal, interpersonal, and classroom and instruction factors on individual and school outcomes. The dominant method used is multilevel, longitudinal cohort research, plus small-scale studies, using a mixed-method approach;
- ‘Learning environments’ focuses on the effects of instruction in the classroom on student achievement. The dominant research method is the design experiment, in which different kinds of learning environments are tested;
- ‘Outcomes of education’ studies the effects of school-based and policy-based innovations. The main research approach is quantitative, using structured questionnaires and existing data sets, as well as controlled field experiments.

Quality

Taking into account the recommendations of the last assessment, a more explicit theory-driven approach has been followed (as clearly reflected in the five key publications), and the programme has gained considerably in coherence and focus. During the past six years, the programme has kept its very good national and international reputation in the domain of school effectiveness, evaluation, and improvement, particularly because of its leading role in the further development of the multi-level approach and its statistical underpinnings. This is reflected by the quality and impact of the publications, several of which appeared in top journals such as *Journal of Educational Psychology*, *Learning and Instruction*, and *Review of Educational Research*. Senior staff members have been invited regularly as keynote speakers at international conferences, and took positions as (co-)editor or editorial board member of international journals or as coordinator of international organisations and as members of national and international scientific review committees.

The programme has been remarkably successful in acquiring research grants and contract research (i.e., €7,500,000), including a high number of competitive and prestigious grants and contracts from NWO and the European Commission and various national funding agencies. Of special importance are TIER (an interuniversity research project aimed at experimental research of education, sponsored by the Ministry of Education and Culture) and COOL5-18

(a large-scale national cohort study sponsored by NWO/PROO, of which it is a leading partner).

In terms of bibliometric outcomes, the programme scores well (including some ‘publication classics’), but not very well, which may to some extent be due to the fact that this is a rather small programme situated in a somewhat specific niche within international educational research.

Productivity

Given the relatively small number of tenured staff in relation to the programme’s large total research staff size, the programme has quite a good academic research output if one takes into account the relatively small number of tenured staff: 157 academic publications (2.27 per research fte), 108 of which are English refereed articles (1.56 per research fte, or 5.46 per *tenured* research fte). During the second half of the assessment period, the number of academic publications increased substantially. The number of professional publications is rather low (96, i.e., 1.36 per research fte).

The number of PhD theses is rather low, namely 13 (6 internal and 7 external; 0.66 per fte for tenured staff). It is problematically low in the first half of the assessment period, but there is a significant positive change during the second half. Clearly, the programme leaders and staff have taken the concerns and recommendations from the previous assessment reports seriously.

Societal relevance

The programme is highly visible and valued by educational policymakers and practitioners, particularly nationally, but also internationally. Programme members have very intensive and productive interactions with important national stakeholders, such as the Ministry of Education and Culture and the Dutch Education Council, and several large public and private school boards. They regularly advise these stakeholders, as they can rely on their very rich and well-managed data sets. Programme members developed several successful interventions for teachers with respect to self-regulation skills, class management, and outcome-oriented teaching, which are being used in the national educational practice. The programme leaders appear regularly in the media. The large amount of acquired contract research is another sign of the programme’s very good societal relevance.

Viability, feasibility and vision for the future

The programme’s viability is considered as very good because of its intrinsic methodological and data-analytical strengths, its leading position in the national and international research on school effectiveness and improvement, its very good scientific infrastructure and resources, its great funding capacity, its central involvement in important new longitudinal data collection and research initiatives such as TIER and COOL⁵⁻¹⁸, and the prominent role of its leaders on the national educational policy and research scene. Plus, the programme may profit from the increasing societal interest for educational effectiveness and evidence-based education. Moreover, the programme leadership has taken into account the criticisms and recommendations of the previous research assessment, which have resulted in further deepening, harmonisation, and focusing of the research programme, in an increase in the scope and scientific reputation of the academic publication outlets, and also in an increase in PhD theses in the past few years.

On the other hand, the programme is confronted with some threats that require serious attention, especially the imbalance between the few tenured and increasingly many non-

tenured and PhD staff members, which may lead to a relatively low availability and work overload of the former. Another point of attention remains the relatively low number of PhD theses, although recently there is clear evidence of improvement.

Conclusion

This is a very good research programme, which is well known in the academic and the educational world, and both nationally and internationally, for its long-standing and methodologically sound research programme on school effectiveness and evidence-based education. Striving for more PhDs, new EC grants, more high-quality academic and professional publications, and continued access to large data sets which are needed for multi-level educational effectiveness research are the main strategic goals for the future.

Programme RUG3: Developmental and Behavioural Disorders in Education and Care: Assessment and Intervention

Programme director: Prof. Hans Grietens, Prof. Erik Knorth, Prof. Alexander Minnaert, Prof. Carla Vlaskamp

Research staff 2011: 6.0 fte tenured, 18.2 total fte

Assessments:	Quality:	4
	Productivity:	3.5
	Relevance:	4
	Viability:	3.5

Brief description

This programme covers three areas: Care for children and young people with (1) learning and educational problems, (2) disabilities, and (3) emotional and behavioural problems. The focus in all areas is on individuals, families or groups manifesting severe problems. Indices of severity are the intensity, complexity, pervasiveness, chronicity or irreversibility of problems and disorders.

The programme aims to:

- Study risk factors and mechanisms explaining problems in children and families related to learning, development and parenting, and protective factors reducing these problems;
- Develop and evaluate instruments/procedures to assess these problems, risk or protective factors;
- Develop and evaluate interventions in the three areas.

Within the discipline ‘orthopedagogics’ the study of these issues implies a contextual approach, which is inspired by ecological and transactional models of human development and well-being. As the aforementioned areas are also studied by other disciplines (e.g., psychology, medicine, law), the researchers of this programme have arranged some collaboration with research groups from different faculties.

Quality

The programme has made some significant contributions to research in its specialist areas, such as the study of outcomes of residential care. The number of publications has steadily improved. However, the quality is mostly average and not exceptional. There are three separate programme areas. It appears that the three areas act autonomously with little interaction between them. Although there was discussion of the programmes being integrated through a focus on early childhood education, evidence of this integration was not presented. There would likely be advantages of cross-programme, cross-disciplinary research. It is encouraging to see that the senior faculty who arrived in the last five years have been successful in getting funding to examine the effectiveness of innovative programmes in the Dutch context and, more generally, to give an impetus to the quality and vitality of the programme. However, there are relatively few publications on the effects of these ‘evidence-based’ programmes.

There is a good track record for obtaining funding for applied, practical and policy-related research, though it is recognised that efforts should be increased to secure grants from national and international competitions.

Productivity

The programme has shown an increase in productivity across academic and professional publications, especially in the second part of the evaluation period (2009-2011). There has been an upsurge in academic research output: 332 academic publications (4.27 per research fte), 123 of which were English refereed articles (1.58 per research fte). While this productivity has shown an increase compared with the period of the previous review, this is a relatively low ratio for academic productivity per fte, particularly given the comparatively large number of tenured staff in relation to the other groups of research staff. On the other hand, the number of professional publications is very high at 394 (i.e., 5.06 per research fte).

The number of PhD theses is rather low – comparatively speaking – at 22 (18 internal and 4 external; 0.69 per fte for tenured staff).

Societal relevance

The programme does substantial contract work with agencies and has strong relationships with professional organisations. Senior faculty have strong reputations and are active in both national and international organisations in their fields of expertise. The programme has produced numerous materials that address the needs of children with disabilities (and their families and teachers), for those working with children in foster care and other residential care context. However, perusal of the programme's website and associated 'urls' provided little information. Given the importance of the web for conveying information, much more could be done as regards social relevance to disseminate its findings.

Viability, feasibility and vision for the future

This programme has undergone considerable change since the previous review. During this time it has been effective in raising its productivity, and this is to a large extent due to the new group of senior faculty that has joined the programme since 2006. Four full professors and one associate professor are nearing or at retirement age, and there is an urgent need to replace them with younger scientists with cutting-edge research ideas in order to introduce a combination of innovation and coherence. However, it was noted that the programme has recently been further strengthened by the recruitment of three new professors: in 2010, one new full professor was recruited and in 2012 two 'professors by special appointment' were hired, each with responsibility for a thematic area of research.

While the programme has relationships with other universities in the Netherlands, Belgium, and other countries in Europe and elsewhere that are working on the same topics, the committee saw relatively little mention or evidence of productive outcomes of this collaboration (except for a few cases).

Conclusion

During the review period this programme has undergone considerable change. The senior faculty members have shown increasing productivity and attracted new funding to do important research. There is a need to use new faculty hires to continue to reshape this programme in order for it to achieve international prominence. More coherence and integration are needed, which could lead to more focused research that will lead to higher quality and productivity.

5. UNIVERSITY OF TWENTE

Programme UT: Educational Design and Evaluation

Programme director: Prof. P.J.C. Sleegers, Prof. A.J.M. de Jong, Prof. C.A.W. Glas

Research staff 2011: 11.3 fte tenured, 28.2 total fte

Assessments:	Quality:	4.5
	Productivity:	4.5
	Relevance:	4.5
	Viability:	4

Since all educational research is conducted within one single programme at Twente University, the committee decided to integrate the assessment at the institute and programme level. Accordingly, the following assessment covers both levels.

1A. The institute

The organisation of educational research at the institutional level is more complex than that typical of other institutes of Dutch educational research. At the University of Twente, researchers are employed in faculties, but since 2011 their research activities are embedded in thematic research institutes. These research institutes provide funding and offer a platform for collaboration to researchers from different disciplines. Like the faculties, the research institutes are positioned directly under the Executive Board of the university. The creation of the research institutes is the result of the UT's general strategic repositioning aimed at creating more excellence, mass, interdisciplinarity, and thematic focus in its research programmes. At the same time, the research institutes put researchers and research teams in an optimal position to establish international alliances and compete for international grants, particularly at the European and international level.

At the faculty level, staff members of the Educational Design and Evaluation research programme are employed in the Faculty of Behavioural Science, where they are housed in three different departments, which are associated with the programme's three different research themes (see 1B). The department heads are also the directors of the research programme.

At the level of the research institutes, the educational research programme is embedded in two interdisciplinary research institutes, both of which include researchers from other faculties and disciplines. Individual researchers of the educational research programme are associated with one of both institutes:

- The Institute for Innovation and Governance Studies (IGS), which conducts research in the fields of social and behavioural sciences, addressing issues of co-ordination, steering and operation of actors and institutions (networks) in both public and private sectors from a multi-level, multi-actor systems perspective. IGS research is organised in five strategic research orientations (SROs), and the educational research programme is involved in two of them: 'Innovation of Governance' and 'Health Assessment and Promotion'. Their main contributions to the latter SRO relate to survey research methodology and computerised adaptive testing.
- The Centre for Telematics and Information Technology (CTIT), which conducts multidisciplinary research in the area of design and application of advanced telematics and information technology systems, and their integration into user environments. CTIT is

purported to be one of the largest ICT research institutes in Europe. Its research is organised in six SROs, and the educational researchers are positioned in the SRO 'Natural Interaction in Computermediated Environments'.

This new organisation reflects management's ambition to enable more interdisciplinary collaboration amongst researchers and is logically related to the interests of the educational researchers of UT (e.g. with some focusing on school organisational and educational assessment issues and others on the uses of educational technology). Indeed, these two research institutes seem to act as stimulating and flexible platforms for the researchers of the educational programme. On the other hand, however, the new organisation means that lines of management and accountability are not very clear to an external observer. Nevertheless, because of the relatively small size of the institution and the informal government style, the relevant managers assured the assessors that this managerial complexity is not problematic in practice. Moreover, most policy-making and research activity still take place at the programme level and at the level of the subthemes or departments (see 1B), rather than at the level of these newly formed interdisciplinary research institutes.

In sum, the leadership functions of and interaction between the different entities seem to work well despite the complex structure. That such a multidisciplinary collaboration has been successfully implemented is, as such, an admirable achievement. However, convincing evidence about whether senior staff from different disciplines is already successfully working and publishing together within IGS or CTIT is not yet available. This reveals how difficult it is to achieve the ambitious goal of multidisciplinary work. Clearly, it takes more than the few years in which the reviewed units have existed.

1B. The programme

The mission of Twente's research programme is 'to develop theories of human learning, instruction, and evaluation and to use these theories to design tools, procedures, interventions, and systems that aim to improve learning in educational and training settings'. As such, the programme is both theory and application oriented, which fits in with the university's 'entrepreneurial' orientation. To realise its mission, the programme uses a multidisciplinary approach with contributions from psychology, sociology, educational science, computer science, human resource management and psychometric theory, and applies a variety of research methods, including experimental studies, survey studies, longitudinal field studies, multiple case studies, and design studies situated in ecologically valid settings. According to the programme's mission statement, it is particularly characterised by its 'educational technology approach'.

The programme distinguishes the following three research themes, which correspond to the three departments of the faculty:

- *Inquiry learning in powerful learning environments* (Department of Instructional Technology; IST) focuses on the investigation of learning processes that underlie inquiry learning and the evaluation of instructional designs for inquiry learning environments. In particular, technology-enhanced inquiry learning environments incorporating computer simulations and/or modelling tools are studied.
- *Educational design and effectiveness* (Department of Educational Organisation and Management; OWK) focuses on the dynamic interplay between curriculum innovation and professional learning in relation to the effectiveness of school and training organisations. Factors that enhance the improvement and sustainability of the quality of

school and training organisations at the system, organisational, programme and individual level are studied.

- *Computerised testing of knowledge and skills* (Department of Research Methodology, Measurement and Data Analysis; OMD) focuses on educational measurement and educational research methodology. Research topics are item response theory (IRT), computerised adaptive testing, optimal test assembly, linking and equating tests, test bias, differential item functioning, and large-scale educational surveys.

However, the above-mentioned programme was established only recently. As a result of the university's new policy and internal developments (e.g. retirements and previous research assessments, productive collaborations between scholars from the different former programmes), it has been decided to integrate the four previously existing programmes (which were evaluated separately in the previous QANU research assessment, RAPES 2007) into a single integrated programme, consisting of the above-mentioned three subthemes. While the first and third subthemes coincide with two of the four former programmes (in terms of scope and leadership), the second programme is a fusion of two previous programmes and has a new coordinator who comes from abroad.

Several steps have been and are being taken towards the integration of the three departments/subthemes. At the programme level, the coordinators of the three departments/subthemes regularly discuss strategic issues and explore future common research plans in informal meetings. Furthermore, there is a very explicit and systematic striving for the mutual exchange of research ideas and collaboration among the three research subthemes/departments. Clear examples of integrative topics are: (1) the development of more adaptive, highly interactive learning environments, and testing of the results in more challenging target groups, (2) ensuring the organisational embedding of these innovative environments in the classroom, and (3) the development and application of advanced structural equation (SEM) models in the context of school effectiveness research and educational surveys. The common responsibility for the new Graduate School (with the leader of the first subtheme as the coordinator) is another important integrating element. However, clear and concrete 'products' of interdepartmental collaboration are still rather rare, and there are no meetings yet for PhD students and staff at the programme level. So, at the moment this new programme does not yet seem as 'experientially real' and important for most of its members.

2. Quality and academic reputation

Overall, the programme is internationally well known and highly respected for its research on technology-based learning environments and on educational measurement and educational research methodology. Other research within the programme on school effectiveness and school improvement is also well recognised to be of an international standard. While particularly the first and also the third well-established subtheme continued to function at a very high level of quality and visibility, the second, newly established subtheme (which integrates two older programmes and elements of the research programme of the new subtheme coordinator) seems still on its way to establishing a clear and strong identity within the domain(s) of educational design and effectiveness. Moreover, the relationship of this subtheme with the claimed integrative theme of the new programme, namely 'information and communication technology', remains somewhat obscure.

The programme's strong international reputation is evidenced by the quality and reputation of the outlets in which the five key publications have been published (including a paper in *Science*). Furthermore, the quality of the general research output is very high, with numerous

publications in leading journals (such as *Learning and Instruction*, *Computers in Human Behavior*, *Contemporary Educational Psychology*, *American Educational Research Journal*, *Psychometrika*). Moreover, the programme leaders and other staff members have been invited regularly to contribute to several prestigious international handbooks, to give plenary lectures at important international conferences, and to become members of editorial boards and visiting scholarships. Particularly the first department/subtheme has received several additional recognitions of scientific excellence, as well as several prestigious prizes for software development.

Publications of the research group are regularly cited by other researchers, as was evidenced, among others, by the programme's very good scores on Prins' bibliometric analysis.

In terms of the nature of the publications, the strategic goals emphasise multidisciplinary and an orientation to the application of results more explicitly than in most other assessed institutes. Both are very challenging goals, which can be typically achieved only after long collaboration between highly motivated staff, both members of the various departments and other scientists of IGS and STTT. However, there has not been enough time to establish sufficiently deep collaboration and valorise it into high-quality, multi-disciplinary publications. Still, there is already some emerging evidence of productive cooperation with other partners of the interdisciplinary research institutes. This is reflected in scientific output in the most recent years in new areas beyond the programme's own specific educational research areas, such as medical sciences, social psychology, technology, and administration. The committee encourages the institute and programme leaders of the UT to carefully evaluate the opportunities and drawbacks of their strategic orientation towards multidisciplinary and practical applications, given that multi-disciplinary and practically oriented research is typically harder to get published in high-quality journals.

3. Resources

Over the review period, staff numbers at Twente's educational research programme have oscillated between 59 and 71 people. However, the number of research fte's has increased from 24.9 in 2006 to 28.2 in 2011. The additional research fte's largely arise through the appointment of more non-tenured staff.

The increase in total research fte's was paralleled by an increase in total funding, from €5,908,000 in 2006 to €6,484,000 in 2011. The ratio between direct funding, research grants and contract research has somewhat shifted over the review period. The share of both direct university funding and research grants in the programme's annual budget has declined slightly. Contract research, on the other hand, is responsible for increasing percentages of the annual budget: from about 22% in 2006 to about 31% in 2011. There are, however, significant differences between the three research subthemes in this distribution. In total, the programme received approximately €13,900,640 in external research grants and contract research over the review period. This makes for an earning capacity of €205,115 per tenured research fte per year, which is high in both absolute and relative terms.

There is a good balance between fundamental and more policy- and practice-oriented grants, and between national and international funds. Remarkably, compared to most other institutes, the amount of total funding per year, and of indirect funding in particular, has remained rather stable over the years. Also, the percentage of direct funding is quite high, compared to other institutes and programmes.

Examples of external funds that have been acquired by the programme during the review period include grants from NWO (One Vidi and one Rubicon grant, 6 PhD projects, 3 post-doc projects and 5 review studies), the EC, OECD, UNESCO, Spencer Foundation, Ministry of Education, Culture and Science (for three so-called Educational Evidence (Onderwijsbewijs) projects and research on data-driven feedback in teams and school performance feedback), SURF, Kennisnet, LSAC (Law School Admission Council, Princeton, USA). A unique feature of the programme is its long-standing success with EU-funded projects, particularly in subtheme 1 (e.g. Kaleidoscope, APOSDLE, CoReflect, SCY, Go-Lab) and in conducting international comparative studies in subtheme 2 (TIMSS, PISA, TALIS).

From a policy perspective, most direct funding resources are managed at the level of the faculty and/or the three departments, with some strategic funding for interdisciplinary research being managed and allocated through the two institutes. This system appears to have been successful in stimulating and priming multidisciplinary activity. Multidisciplinary applications for external research funding are also encouraged. Because of wider economic factors, the levels of direct funding by the university have not increased in recent years, but levels of resources still seem adequate. The importance of finding other sources of income for the viability of the research programme is rightly identified as a challenge in the programme's SWOT analysis.

Given the programme's central technology-oriented mission, there is an acknowledged need to be adequately funded for investigating and pursuing all the imaginable opportunities that new technology is increasingly offering. In principle, given its current and future research agenda as well as its past successful experiences with EU-funded projects, it should be relatively easy to make continued use of the opportunities offered by the EU funding programmes which are generously offering support to research in the areas of interest in Twente.

4. Productivity

Research productivity of staff is the joint responsibility of the leadership of the faculty (and its departments) and of the two interdisciplinary research institutes. Criteria for staff tenure and promotion decisions include the usual factors of productivity in achieving publications, completion rate of doctoral students as well as gaining external funding. Overall, institutional productivity is quite high.

During the review period the research staff of the programme produced 620 academic publications. The total academic output of the programme has increased over the years: from 90 academic publications in 2006 to 112 academic publications in 2011. On average, the programme produced 3.8 academic publications per research fte per year during the last six years. While the input-output ratio for academic publications in general has remained quite stable during the review period, the production of English refereed articles has significantly increased over the years: from 47 English refereed articles in 2006 to 74 in 2011. In total, the programme has produced 370 English refereed articles over the review period, which equals an average of 2.2 English refereed articles per research fte yearly. These are very good productivity scores for academic output.

In total, the programme has produced 458 professional publications and products over the review period, which equals a yearly average of 2.8 professional publications and products per research fte. Over the review period, the absolute number of professional publications and products has remained quite stable, while the production of professional publications and products per research fte slowly decreased: from 3.1 professional publications and products

per fte in 2006 to 2.4 in 2011. (However, as the lists of professional publications also include scientific conference papers, these numbers are difficult to interpret.) This slow decrease may be related to the publication policy at the institute and department level to publish more in academic outlets and more specifically in English journal articles.

In total, over the review period 2006-2011, the programme produced 60 PhD theses, 14 of which were external ones, while 46 were internal ones. This means that there was an average of 0.9 PhD theses per tenured fte, which is quite good.

In principle, it would have been interesting to analyse in detail whether the goal of interdisciplinary cooperation amongst researchers is being realised in the corresponding production of interdisciplinary publications. However, the available documentation did not allow such a detailed analysis to be carried out.

5. Societal relevance

The societal relevance of the work of educational researchers at the University of Twente over the assessment period is very clear. University and faculty managers consider that this is being enhanced by the university's 'entrepreneurial' philosophy and by the encouragement and priming of interdisciplinary research.

Probably the most manifest and continuing strength of the programme in this respect is the production and dissemination by subtheme 1 of a set of widely used, computer-supported, learning environments (which are distributed commercially by national and international publishers). For instance, currently 70,000 licenses of the ZAP software have been sold worldwide. The SimQuest software has led to one spin-off company. The different international prizes won for these products further testify to their quality and societal impact. Recently, subtheme 3 has initiated similar initiatives in the area of statistical software.

The societal relevance of the programme is also expressed through the intensive consultancy activities of the programme director, who is responsible for subtheme 2.

Staff members participate in relevant local, national and international consortia with stakeholders from industry, non-governmental organisations and government. Some of them are particularly active at the international level (e.g. PISA and ESS).

Staff members, particularly of subtheme 3, participate in the Research Centre for Examination and Certification (RCEC), which was established in 2009 in collaboration with CITO (national institute for educational measurement). RCEC carries out externally funded PhD projects aimed at educational assessment. Together with a large number of external PhD students from third-world countries in subtheme 2, these RCEC PhD students are responsible for the very high number of external PhD candidates (i.e., 55) currently affiliated with the research programme.

The programme's research results are disseminated outside the scientific community through various newspapers, magazine articles, and interviews for radio and television. Also, staff members are on the editorial boards of various professional journals.

6. Strategy for the future

According to the self-evaluation report, several factors contribute to the viability of the research programme.

Firstly, the programme has not only been successful in continuing its extensive NWO, EU and OECD funding, it has recently also successfully explored new sources of funding. Following the successes with prestigious Vidi and Rubicon grants, the programme leaders plan to put more effort into obtaining more prestigious grants. In addition, the major themes of innovative educational uses of digital technology, the analysis of school effectiveness at the international level, and educational measurement and the various new topics at the intersection of these areas of research offer good prospects for continued external funding, also at the European level.

Secondly, the appointment of a new professor in the field of leadership and school improvement, who also became programme director in 2008, acts as a unifying force for the newly established programme. The above-mentioned examples of emergent cooperation seem to promise that this recent merging operation will bear fruit, but it is at the moment too early to judge whether it actually will.

Furthermore, the self-evaluation report states that the programme holds a strong strategic position within the University of Twente as a result of the university's new strategy to create more focus and mass in its educational and research programmes (cf. section 1A). Clearly, the fact that the programme is strongly embedded in two research institutes can be considered to strengthen the programme's prospects for the future, particularly given its traditionally strong reliance on EC grants, where critical mass and interdisciplinarity are important elements for success. However, it remains to be seen if the programme's somewhat 'hybrid' relation with these two – quite different – interdisciplinary research institutes will contribute to the intended unifying process at the programme level. More concretely, one may ask why the two corresponding SROs have to exist in two different research organisations without a more explicit linkage between them. Just to give an example: does digital technology (as developed and studied within CTIT) not only make possible the creation of dynamic learning environments, but also offer new tools for use in adaptive testing (as developed and studied within IGS)?

The SWOT analysis further mentions the programme's national and international visibility and reputation as an additional strength. This certainly holds true for the well-established first and third subthemes, as well as for the new second one, which also contains elements of strong and international visibility (for instance, in the domains of international comparative studies of school outcomes and educational innovation).

On the negative side, the newly started local Graduate School, 'Learning and Educational and Training Settings', does not attract many PhD students (from abroad) and is still in the process of providing a strong graduate programme for future PhD students (cf. section 7). The further development of this graduate programme is therefore an important point of attention, as acknowledged in the institute's self-assessment report.

Second, many studies are carried out in realistic (school) contexts and involve curriculum-related instructional and testing materials. This jeopardizes quick and easy success with respect to productivity, as the design of such studies requires much time and energy. However, the committee would not want this consideration to stop the UT working in realistic (school) settings.

A third possible threat to the programme's future relates to the coordination of EU projects. Although the programme has greatly profited in many respects from the funding and opportunities offered by these EU projects, the coordination of such large-scale projects is

very time-consuming and does not in itself result in scientific output in the form of publications and PhD theses. Finding a good balance between the costs and benefits of EU projects is another challenge for the future.

Finally, the general goals emphasising the interdepartmental and the multidisciplinary work of researchers of education represent good preparation for the future. However, the specification of future plans was not as strong as one would hope. There seems to be a lack of a clear vision on how to link the second research theme with the programme's major technologically oriented research focus. Furthermore, at a more specific level, an explicit vision of where education is going through the explosion in the use of tablet computers in school environments seems absent. It would be a missed opportunity if this issue were not included in the strategies for the near future – in terms of the recruitment and funding policies associated with research.

7. PhD training and supervision

The formal training of PhD students takes place in the ICO or IOPS national research schools. In addition, the university created a research school for the organisation of graduate education as a whole two years ago. This research school has a special programme for the educational PhD students (entitled: 'Learning in Educational and Training Settings') that is complementary to the national research schools. As stated in the self-assessment report, this new programme is still in its infancy and is therefore not yet attractive to good PhD students from abroad. No research master programme has been available, but one is planned for the future, possibly in cooperation with another university. This might also help to attract good PhD students.

The work activities of PhD students are organised in various ways, depending on the department (i.e. research subtheme) with which they are associated. There are no organised meetings for PhD students at the programme level (which may be one of the reasons why PhD students do not seem to identify themselves already with the new research programme), but they are involved in relevant meetings within the institute of which their supervisor is a member. The PhD students interviewed were very positive about these arrangements, stating that they had ample opportunities for mutual contacts, enough regular contact with supervisors, and that links with relevant researchers in other universities were actively encouraged and facilitated, especially via the national research schools. Students seemed to highly appreciate this involvement at the national level. Overall, the system seems to work effectively in supporting and developing postgraduate work, though some students would appreciate more training and guidance in the local PhD programme in seeking job opportunities outside academia, as well as in academic writing and in the ethical aspects of research.

Out of the 23 PhD students who enrolled between 2003 and 2007, 48% had graduated after four years. After five years another 26% had graduated, and another 9% after six years. Only 4% have not yet finished their projects, while 9% of the total of 23 students dropped out. Thus, PhD efficiency is comparatively very high.

This success may to some extent be explained by the good tracking system, which allows evaluation of student progress and the extent and quality of their research outputs. All PhD students start their research projects with an education and supervision plan. This plan outlines the scientific assignments, the PhD training activities, and the planning of supervision and is evaluated on a yearly basis, with a go/no go decision after the first year.

On the other hand, the highly multidisciplinary environment in which they have to operate may create extra problems for PhD students. For example, the responses of staff to student needs may be slower than in more typical research environments, particularly for PhD students coming from abroad. It is also likely that the interests of some domains – such as ICT in the CTTT – may dominate and make it difficult for the representatives of educational interests to organise PhD studies efficiently in the crossfire coming from different directions.

8. Conclusion

Reading the documentation and listening to the interviews with the representatives of the education research of Twente lead to the conclusion that the structure of the organisation of educational research is complex. However, individuals with appropriate expertise can flexibly move between entities, orientations and themes to multidisciplinary environments that best fit the specific research problem associated with education. Twente's approach can be recognised as an original and promising way in which to organise research, so as to respond to the real-life questions which seldom can be answered through traditional, mono-disciplinary approaches

At the programme level, Educational Design and Evaluation is a programme with generally very high quality, productivity and societal relevance, and with several elements of international excellence, particularly in its first subtheme. However, while the viability of the three distinct parts also seems very good, it remains to be seen if they will actually develop into an integrated programme in terms of research goals, outcomes, policy and culture. The further development of the recently established local PhD programme and the installation of a high-quality and attractive research master seem to be the major challenges for the future.

6. UTRECHT UNIVERSITY

A. Institute level

1. The institute

The Pedagogics and Educational Science research programmes of Utrecht University are part of the Faculty of Social and Behavioural Sciences (FSBS) and are embedded in the faculty's Research Institute of Social and Behavioural Sciences. The research is subdivided into two large programmes, which are associated with a corresponding faculty research cluster:

- Child and Adolescent Studies (CAS, UU1), which consists of two research groups from Pedagogics (Development and Treatment of Psychosocial Problems, and Adolescent Development: Determinants and Characteristics) and one from Interdisciplinary Social Science (Youth in Changing Cultural Contexts);
- Education and Learning (EL, UU2), consisting of research groups from Pedagogics (Development and Education of Children with (Mild) Cognitive and Motor Disabilities, and Learning in Interaction), and the former IVLOS, the university's centre of Education and Learning.

The research groups that together constitute a programme are organised around one or two full professors and their tenured staff. The research group level is the level at which the most intensive daily research activity takes place. Scientific meetings are organised at this level, but also at the level of the programme. Important decisions with respect to the mission, structure, and strategy of the programme are taken at that level. Each programme is led by its own board of full professors, one of whom acts as chair. The move towards larger entities (programmes) was motivated by strategic considerations, particularly to create more synergy among researchers and to compete for indirect grants at the university, national and EC level.

To encourage multidisciplinary research, Utrecht University defined fifteen multidisciplinary research clusters in 2005, including research groups from different faculties, the so-called research focus areas. The CAS and EL programmes participate in the research focus area 'Coordinating Societal Change: Life Course Dynamics, Economic Flexibility and Social Cohesion' (CSC), together with research groups from the Faculty of Law, Economics and Governance and the Faculty of Geosciences.

In 2011, a new research impulse area, namely Educational and Learning Sciences (ELS), was added to the research focus areas. The EL programme is the core of this research impulse area, and it collaborates with groups from the Faculty of Humanities, Faculty of Sciences, Faculty of Geosciences and the Faculty of Medicine.

Besides the research focus areas, Utrecht University has recently developed four 'strategic themes' for future research, following its Strategic Plan. The CAS and EL programmes are associated with the 'Youth and Identity' theme, in which the research focus area CSC as well as the research impulse area ELS participate. The CAS programme constitutes one of the core research groups in this theme. The Faculty of Social and Behavioral Sciences (FSBS) initiated the 'Youth and Identity' theme and acts as its commissioner. Furthermore, FSBS, in cooperation with the University Medical Centre Utrecht, has established the Youth Centre, which is closely related to the 'Youth and Identity' theme. The Youth Centre will study risk and resilience over the early life course and will coordinate data collection on biological, neuropsychological, family and societal factors of a cohort of infants followed into childhood

and a cohort of school-aged children followed into adolescence. Both the CAS and the EL programme will play an important role in the Youth Centre's research activities.

2. Policy on scientific quality and academic reputation

The quality and academic reputation of the scientific work of the two programmes is very high, according to international standards, as further documented in the programme sections of the self-evaluation report. In fact, the strong commitment to an interdisciplinary research approach within the faculty along with the interdisciplinary university initiatives seems to have a positive impact on the research quality and productivity. The collaboration and mutual exchange between the two research programmes seem to be rather modest, however, but it is considered understandable given the size of the programmes.

The institute section of the self-evaluation report mentions several general indicators of the academic quality of the programmes. These comprise the more 'classical' ones (i.e., numerous editorships and memberships on editorial boards of international journals, participation in advisory and review committees, fellowships, keynote lectures at (inter)national conferences, the organisation of conferences and the coordination of PhD education activities within the national research schools), as well as the prizes received by young scholars and the staff members' success in obtaining research grants and NWO funding. More details are provided in the programme evaluations.

Since one of the two programmes (namely CAS) decided not to participate in Prins' bibliometrical analysis, this quality aspect will be addressed at the programme level of EL.

3. Resources and resource policy

Staff numbers at the Pedagogics and Educational Science research programmes have gradually increased over the review period: from 120 persons (equal to 40.4 research fte) in 2006 to 137 persons (equal to 54.4 research fte) in 2011. CAS staff has grown more than EL staff. Significantly, not all categories of staff have increased. The increase of staff primarily concerned the PhD students: from 41 PhD students (equal to 17.4 fte) in 2006 to 56 PhD students (equal to 27.7 fte) in 2011.

The increase in research fte was paralleled by an increase in total funding. Total funding of the institute stood at €3,542,000 in 2006, while it gradually grew to €5,803,000 in 2011. The ratio between direct funding, research grants and contract research has also shifted over the review period. Direct funding by the university has decreased substantially from 58% of the total income in 2006 to 40% in 2011. While the percentage of the total income brought in by contract research remained more or less equal, the income generated by external research grants has increased from 32% of the total income in 2006 to 43% in 2011. In total, the programmes received €15,014,830 in external grants over the review period. This makes for an earning capacity of €178,726.70 per tenured research fte per year, which is very high, comparatively speaking.

The share of the total funding attributed to the two research programmes gradually shifted over the review period. While CAS received 44% of the annual budget in 2006, it received 54% in 2011. The EL programme, on the other hand, received 56% of the annual budget in 2006, and 46% in 2011. This change reflects the fact that the total budget of both programmes has grown significantly, but even more so for CAS than for EL.

The generally decreasing chances of obtaining NWO funding necessitate finding other sources of income for the viability of both research programmes. In this respect, two

promising factors have been identified. First, the model used by the university to allocate funding among faculties just changed in favour of the Social Sciences. In fact, first-stream money is now influenced by (1) the number of students, (2) the success in gaining second-stream funding and (3) the number of completed PhDs (third-stream contract research funding is explicitly left aside, because of its varying scientific quality and lack of quality control). The higher the amount of second-stream funding and the number of completed PhDs, the more first-stream money is allocated. Secondly, the research policy of Utrecht University and of FSBS in particular is considered its strength. Both the research focus area CSC and the research impulse area ELS benefit from substantial extra funding from the university as well as FSBS. CAS and EL, being core participants in CSC and ELS, will benefit from this extra funding. The same is the case for the 'Youth and Identity' strategic theme, in which CAS and EL play significant roles. The self-evaluation report states that the theme can count on considerable new investments from the university and on additional funding from the participating faculties and FSBS in particular. Moreover, because of its strong positioning in these larger interdisciplinary research impulse areas and strategic themes, the two programmes are well positioned to compete on the EU level, where a new framework programme for research and innovation has been developed ('Horizon 2020').

4. Productivity strategy

Over the last six years, the research staff of the joint programmes produced 1285 academic publications. During the review period the total academic output of the joint programmes shows a rather versatile image, with numbers ranging between 201 and 233 publications per year. Moreover, the total output does not demonstrate a clear line over the years, although it should be noted that the total number of publications in 2011 (233 publications) is the highest peak in the review period.

Although the total academic output is versatile, the input-output ratio is rather stable. The institute produced on average 4.3 to 4.4 academic publications per research fte yearly. An exception to this was the year 2006, during which on average 5.4 academic publications per research fte were produced. Paying attention in particular to the number of English refereed articles, it is noted that over the last six years the two programmes jointly published 789 English refereed articles. The production per research fte increased over the review period and had a yearly average of 2.8 English refereed articles per research fte, which can be considered as very high.

The programmes were not equally productive, with CAS being somewhat more productive than EL, both for academic publications in general and for the English refereed articles, and both in absolute terms and in relation to the available research fte per year. In contrast, both in absolute and relative terms, the number of professional publications was somewhat higher for EL than for CAS. The two programmes jointly produced 363 professional publications and products over the review period. The production slowly decreased over the years and shows a moderate yearly average of 1.3 professional publications per research fte.

In total, over the period 2003-2007, the programmes together produced 108 PhD theses, 22 of which were external ones. This means that there was an average of 1.3 PhD theses per tenured fte, which is very high, relatively speaking.

5. Societal relevance

The societal relevance of both programmes seems to follow rather directly from their missions, which consist, for CAS, of contributing to the quality of the professional field by developing, testing and evaluating interventions for the prevention, diagnosis and treatment

of psychosocial problems in childhood and adolescence. As for EL, the self-evaluation report states that the programme connects to many of the crucial components of the national and European policy agendas on education. In these policies, education is considered to be the key to sustainable social and economic development and improving the quality and efficiency of education, levelling the playing field for more equity, preventing school drop-out, fostering excellence, innovating workplace learning and raising the competences of the teaching force are core priorities. The EL programme reportedly contributes to many of these priorities.

Both programmes realise the potential impact of their scientific work, by actually implementing the research-based pedagogical and educational interventions and developing and distributing other kinds of practically relevant products of their research (diagnostic tools, observation schemes, assessment tools) in close collaboration with various local and national institutes and organisations in the corresponding professional fields.

Moreover, programme members are active as consultants of national and international practice and policy making, and serve on policy advisory boards of the Dutch and European government, and on editorial boards of practically oriented journals, fulfil consultation roles for municipalities, educational organisations and parent associations, and speak at conferences for practitioners, professionals and policymakers.

Finally, sharing of knowledge takes place by, for instance, disseminating research results to the general public and giving interviews for newspapers, radio and television.

With respect to policy, the university is working on an inventory, categorisation and assessment system for the societal relevance of its programmes and research groups, as an additional sign of the value it attaches to this societal aspect of the programmes' mission. However, at the moment, there are no plans to include elements related to societal impact (such as contract research grant acquisition) in its allocation model (cf. section 3).

6. Strategy for the future

The committee was impressed by the analysis of the research environment elaborated in the self-assessment report, in which the two programmes have to operate at the EU, national, university, and faculty levels. It was equally impressed by the well-articulated and dynamic research policy of the faculty government in this respect.

While the viability of the research programmes CAS and EL primarily depend on their past performance and intrinsic merits, it seems clear that their embedding in a university and faculty with an ambitious and future-oriented research policy can be considered an extra strength.

The faculty and programme leadership define the strategy for the future in terms of maintaining and strengthening the position of the programmes within FSBS and Utrecht University. Moreover, the participation in the research focus area CSC and the 'Youth and Identity' strategic theme ensures their position at the university level. In this respect, considerable investments have been made (and will continue to be made) by the faculty with the specific aim to position the two programmes as strongly as possible within this broader dynamic research area. It can be expected that the programmes will continue to benefit from these extra investments by the FSBS and the university in the research focus areas and the strategic themes. Within this context, it is also considered favourable that the research focus areas and strategic themes show characteristics that are well in line with the Dutch government's policy and the EU Horizon 2020 programme.

As it will be harder to obtain NWO funding due to the Dutch government's top sector policy, according to the self-evaluation report, FSBS will invest substantively in its Grant Office, thus providing further improved support in preparing grant applications, particularly for applications in EU funding schemes.

Another recent but very important indicator for the viability of the two programmes is the result of the 2012 NWO Gravitation Programme competition, in which the Individual Development application, a consortium of research groups at Utrecht University and other Dutch universities coordinated by FSBS and CAS staff members and with core participants from CAS, received substantial funding. It is expected that Individual Development will play an important role in research in the field of Pedagogics and Educational Sciences at FSBS and Utrecht University in the upcoming years.

Finally, the committee took note of the fact that, as a consequence of its very good performance in the allocation system for the university's direct funding (based on their recent accomplishments in terms of student numbers, acquisition of second stream research funding, and completed PhD theses), the direct funding of FSBS will increase in the near future, which will further help it to face future challenges.

7. PhD training and supervision

PhD training is organised both at the faculty level, in PhD training programmes, and at the national level, in research schools. At the faculty level, both CAS and EL maintain PhD training programmes tailored to their specific needs. These PhD training programmes are integrated in the Faculty's Graduate School of Social and Behavioural Sciences of FSBS, which offers general academic training and courses for its PhD candidates. According to some PhD students, more attention could be paid to research ethics.

At the national level, the ISED and ICO research schools organise specific courses and meetings for PhD students and support interuniversity networks of PhD students for the two programmes. The PhD students appreciate that these networks allow them to follow advanced courses at other universities, and bring them in contact with other PhD students from other institutes and programmes with other theoretical and/or methodological approaches.

The institute and programme leadership is pleased with the way in which the PhD training is shared between national and local doctoral schools, and is strongly against the pressure to restrict the PhD finalisation term to three years. Both PhD programmes also build upon research master programmes: Development and Socialization in Childhood and Adolescence (DaSCA), which is maintained by CAS, and Educational Sciences (EdSci), which is maintained by EL. These two-year international research master programmes provide opportunities for young talented students to move on to the level of PhD students.

Out of the 44 PhD students who enrolled between 2003 and 2007, 25% had graduated after four years. After five years another 43% had graduated, and after six years another 14%. Thus, after six years, 82% of PhD students had finished their projects, while 16% was still working on them, and 2% had dropped out. These data are very good.

Most research groups of the two programmes have meetings with the research staff every two weeks. There are also monthly meetings at the programme level. Apart from these more formal meetings, there are frequent meetings with the supervisors.

According to the PhD students, most supervisors create enough time for these meetings, and students do not have to wait long for an appointment. Most PhD students have two or even more supervisors, including a daily supervisor who is very accessible.

Remarkably, all PhD students do teaching and/or provide teaching assistance for at least 10% of their contractual time. PhD students enjoy and appreciate this part of their job very much.

B. Programme level

The committee comes to the following overall programme scores for quality (Q), productivity (P), relevance (R), and vitality/feasibility (V) for the two programmes of the Faculty of Social and Behavioural Sciences of Utrecht University:

Code	Programme name	Q	P	R	V
UU1	Child and Adolescent Studies (CAS)	4.5	5	4	5
UU2	Education and Learning (EL)	4.5	4.5	4.5	4.5

The detailed assessment for each programme follows in the next section of this report.

Programme UU1: Child and Adolescent Studies

Programme director: Prof. Maja Dekovic (chair)

Research staff 2011: 7.0 tenured fte, 28.6 total fte

Assessments:	Quality:	4.5
	Productivity:	5
	Relevance:	4
	Viability:	5

Brief description

This programme is dedicated to the study of child and adolescent development. Its overarching research aims are to explain how individual characteristics, proximal social relationships, and the wider social and cultural context shape developmental trajectories and improve the prevention and treatment of negative developmental outcomes. These aims are addressed in a multidisciplinary way, through the collaboration between research groups originating in Pedagogics, Interdisciplinary Social Sciences, and Developmental Psychology (the latter is excluded from the current review). CAS has its own PhD training programme within the faculty's Graduate School and related research master. The CAS research programme is part of the Faculty of Social and Behavioral Sciences and is one of the key players in Utrecht University's strategic theme 'Youth and Identity'.

Quality

The programme has a clear common goal and mission, as well as common research methods comprising longitudinal studies and experimental designs. The cooperation of several disciplines in one programme has clearly been fruitful in terms of the breadth and depth of the research agenda, the attraction and education of young scholars, and the collaborations with other research programmes (inter)nationally. The quality of the research is reflected in the high number of articles in top journals in the field, such as *Child Development*, *Journal of Child Psychology and Psychiatry* and *Development and Psychopathology*. The programme leaders have taken positions as associate editors or editorial board members of respected international journals in the field of child and adolescent studies, indicative of their increasing international recognition and prominence. CAS members have been very successful in earning competitive and prestigious grants (e.g., NWO Innovational Research Incentives Scheme). They occupy key positions in the university's strategic theme 'Youth and Identity' and are a prominent partner in interdisciplinary collaborations (e.g., Youth Center), while maintaining a strong socio-behavioral focus. The programme is dynamic and growing and offers a stimulating and rewarding environment for young scholars. Its success in educating PhD students is reflected in the awards they receive for their papers and dissertations and their success in attracting prestigious grants and pursuing academic careers.

As for the programme's earning capacity, it is stated that during the review period CAS acquired grants and contracts worth €7,165,360 euro, leading to an amount of €186,452 euro per tenured fte. These are very good figures. Over the review period, total funding has doubled (from €1,557,000 euro in 2006 to €3,023,000 euro in 2011). This is primarily due to the strong increase in research grants: percentages of funding through research grants more than doubled from 21% to 54% of the total budget.

Productivity

The programme has an excellent academic research output: 729 academic publications (5.1 per research fte), 450 of which are English refereed articles (3.2 per research fte). The large number of Dutch refereed articles also evidences the success of the programme's broad, nationally and internationally oriented publication strategy during the evaluation period (88 in total or 0.6 per research fte). On the other hand, the number of professional publications and products is quite modest: 168 (1.2 per research fte).

During the evaluation period 41 PhD theses were completed (29 internal and 12 external), equalling 1.1 per fte for tenured staff. In comparative terms, this is a very good PhD productivity score.

All of these indicators underline the programme's excellence in terms of productivity.

Societal relevance

The programme contributes to the quality of the professional field by developing, testing, and evaluating interventions for the prevention, diagnosis and treatment of psychosocial problems in childhood and adolescence. CAS members collaborate with many institutes in the professional field (e.g., young mental health care institutions, hospitals, youth care). Their output is also indicative of a strong commitment to the dissemination of findings to professionals as well as the general public. By participating on accreditation panels and advisory boards, they contribute to the scientific quality of professional work in the Netherlands. Programme leaders have also contributed significantly to the development and implementation of effective interventions and the national policy regarding youth care.

Overall, the societal quality and impact of the programme are considered very good, although there is still room to improve its societal impact and visibility.

Viability, feasibility and vision for the future

The programme's viability is considered excellent. CAS has been the initiator of the strategic theme 'Youth and Identity' at Utrecht University, and its collaboration with other leading programmes at the university is expected to strengthen its position and impact even more. Also the Youth Center has been recently established, in collaboration with medical sciences. This opens opportunities for conducting innovative research on the impact of brain development on developmental trajectories. CAS members have acquired extensive external funding in recent years, and CAS is participating in the large-scale NWO Gravitation Programme 'Individual Development' (2013-2022; 27.6 million euros). This new grant will provide substantial additional funding in the next few years for new longitudinal data-collection. The high-quality PhD programme and the stimulating and positive research climate make this programme very attractive for the next generation of researchers.

Conclusion

Overall, this is an internationally competitive research programme, with a recognisable common goal and mission. The number of PhD students and funding by grants has grown substantially in the past few years. The programme's productivity is excellent, and the output covers a broad range of categories. The strategic decision to raise the number of high-impact publications has been successful. Substantial additional funding in the coming years and a strong collaboration within the programme and with other disciplines will likely foster its international reputation and innovative power. Although the programme clearly contributes to the quality of the professional field, an explicit strategy to increase societal relevance will further strengthen the programme.

Programme UU2: Education and Learning

Programme director: Prof. Paul Leseman

Research staff 2011: 9.0 tenured fte, 25.8 total fte

Assessments:	Quality:	4.5
	Productivity:	4.5
	Relevance:	4.5
	Viability:	4.5

Brief description

According to the self-evaluation report, the Education and Learning programme studies education and learning in an integrated framework, focusing on the regulation of learning processes by neurocognitive, affective, physical, social and cultural factors, in both informal and formal education settings, during key phases of life.

The programme's central idea is that regulating learners' attention and action through modelling and scaffolding, and eliciting metacognitive self-reflection and explanation to others, are essential to move learning beyond the immediate situation and to integrate learners' intuitive understanding with conventional verbal-symbolic knowledge.

The programme consists of three subgroups. 'Early childhood' studies the development of learning-related skills in the micro-contexts of the family and preschool in relation to learning in formal education. 'Primary, secondary and tertiary formal education' studies how learning is regulated in interactions involving teacher-to-student, peer-to-peer, student-to-subject matter and student-to-learning materials, with attention to both the instructional and the socio-emotional interactions. In 'work life', vocational and professional expertise development in the dual context of workplace and formal learning is studied.

Quality

Even though the programme is very large in terms of size and scope and combines four former programmes, it has a strong internal synergy, which is evidenced by the shared conceptual framework and shared research methodologies and by joint research activities (e.g., seminars at the programme level, discussion groups and research projects in which the different subgroups are involved). The programme is very well embedded in the UU research area, but also nationally and internationally.

The programme has contributed to international research through several high-quality publications reporting key findings on topics related to all three parts of the programme. The very good scores in Prins' citation analysis reflect this, among others. However, these scores are to a large extent due to the impact of senior research leaders, who have since left the programme. Members of the tenured staff publish frequently in the widespread journals of their field and regularly also in top journals such as *Review of Educational Research*, *Reading Research Quarterly*, *Learning and Instruction*, and *Computers and Education*. They are regularly invited to give keynote lectures at international conferences, and programme members have acquired several national and international awards and prizes. Their academic reputation is further evidenced by many editorial functions in national and international journals, by leading functions in national and international scientific committees and boards, by the organisation of influential international conferences, and by coordination responsibilities in the national research schools.

The earning capacity of the group is high: it is stated that the programme acquired grants and contracts worth €6,740,770. This is equivalent to €147,888 per tenured fte, which is quite good. Many of the acquired research grants are competitive university grants and national grants (NWO), but the European funding is still rather small. There is a lot of contract research, too.

Productivity

The group has produced a total of 553 academic publications, of which 336 are refereed English journal articles, equivalent to 3.85 academic publications and 2.34 refereed English publications per fte. These are very high productivity scores for academic publications. Interestingly, staff members continue to publish regularly in Dutch scientific outlets, too. The number of professional publications and products is reasonable (193, i.e. 1.34 per fte).

A total of 67 PhD theses was completed (of which 57 were internal). There are strong fluctuations over the years with no clear increasing or decreasing trend, but relatively speaking, the PhD production is also very high, namely 1.47 per tenured staff fte.

Societal relevance

Societal visibility, relevance and service of the programme and its members seem to be strong points of the programme. The EL programme contributes actively and visibly to improving the quality and efficiency of education in The Netherlands (and abroad) through numerous activities and initiatives of societal relevance, in each of its three major subdomains. Staff members develop and make available diagnostic tests, prevention and treatment programmes, and scales for classroom observation at various educational levels. These are intensively used in national (and to some extent international) educational practice. Staff members serve prominently on policy advisory boards of the Dutch and European government, professional organisations and broader societal organisations, give lectures for practitioners, professionals and policymakers, and appear in the media. They also are active in the faculty's educational consultancy and training group and the clinical assessment centre.

Viability, feasibility and vision for the future

This programme is well-embedded in the UU's research area and policy: it constitutes one of the five research priority clusters of the faculty, it is the core of UU's research impulse area Educational and Learning Sciences (since 2011), and it is also involved in UU's interdisciplinary research focus area Coordinating Societal Change. Moreover, it is very well embedded nationally and internationally. Other indicators of the programme's vitality are that already a high number of prestigious new research grants have been acquired in 2012, that several promising new scholars have joined the programme, that new scientific infrastructure has been acquired and that interdisciplinary cooperation with the CAS programme and with other research groups of the UU is emerging.

There are however some important challenges for the future. First, the international prominence and impact of the programme's research output remain an issue of continuing attention, particularly since several of the professors with the highest scientific productivity and reputation have left the programme. Second, and related to the first point, the programme still has to find a new leader for the third subdomain. Third, given serious concerns about cuts in the national funding for education and learning, enhancing the fund-raising capacity at the EC level is a priority for the future.

Conclusion

This is a very good programme in terms of quality, productivity and societal relevance, with a very good perspective for future viability and growth. The integration of the educational research at UU into this larger programme seems to have been very successful, and the programme seems to have properly faced several changes in leadership during the assessment period. Bringing the international impact of the scientific output of the new generation of research leaders and new tenured staff members to a higher level of excellence and successfully applying for European research funding seem to be the major challenges for the future.

7. UNIVERSITY OF AMSTERDAM

A. Institute level

1. The institute

The Research Institute of Child Development and Education (RICDE) is one of four research institutes within the Faculty of Social and Behavioural Sciences (FMG) of the University of Amsterdam (UvA). It was established in 2009, when the Division of Fundamental Research of the SCO-Kohnstamm Institute (SCOKI) and the research programme of the Graduate School of Teaching and Learning (GSTL) were merged into two new, larger programmes:

- Child Development (RPCD, UvA1), which includes empirical research on cognitive, social-emotional, and biological development, and the effects of contextual factors and interventions on child and adolescent development and behaviour (normal and abnormal);
- Education (RPEDU, UvA2), which includes empirical research on learning, learning disorders, and instruction, and the effects of learner characteristics, specific educational interventions and contextual factors on the acquisition and development of cognitive and social-emotional skills.

At the same time, the Division of Applied Research of the former SCOKI was transformed into an independent, commercial company for commissioned research within the UvA Holding BV ('The Kohnstamm Institute').

As part of the university policy, the teacher-training programme has been incorporated into RICDE.

From 2006 onwards, seven full professors have been appointed to replace the five former programme leaders of SCOKI/GSTL. Each of the professors supervises one or more lines of research within the two programmes, and together they form the steering committee that advises the director of RICDE.

The committee notes that during the period under review, the institute has seen a substantial amount of reorganisation. Initially, the lack of continuity of senior management seems to have caused some problems, but the seven recent appointments have remedied this. A more stable management structure has now been established, in which all professors meet once a month to discuss current activities and future plans. This appears to provide a logical and stable basis for the organisation and management of research. As a result, the research programmes have gained in strength, although coherence across the institute remains an issue. In this respect the committee was pleased to learn that all staff, formerly spread over different locations, are currently housed in one building.

2. Policy on scientific quality and academic reputation

To underline the scientific quality and international academic reputation of the research conducted at the institute, the self-evaluation report lists several indicators of academic esteem. Members of the research staff have, for example, served as editors of book series and as editors in chief, associate editors or editorial board members of leading (international) journals. They frequently participated in dissertation committees in the Netherlands and abroad, served as reviewers of research proposals and were members of the boards of (inter)national academic organisations. Also, research staff members delivered keynotes at

various conferences, were awarded international prizes and co-organised various international conferences, symposia and invited colloquia.

The committee further applauds the institutional emphasis on quality of output, rather than on quantity. The self-evaluation report states that although the institute is generally 'pleased' with the quantity and quality of publications, it stimulates researchers to aim for high-impact journals in order to increase their chances of acquiring external funding. Also during the interview, the institute management confirmed that it values quality above quantity, for instance in the selection criteria for tenured positions and promotion.

The quality of the research is underpinned by the outcomes of the citation analysis by Prins. For both RICDE programmes, the impact of their 60 best publications was found to be much larger than the average impact of the leading journals in the various domains. Actually, both RPCD (UvA1) and RPEDU (UvA2) ranked very highly among the pedagogical and educational research programmes, respectively, involved in this assessment. The self-assessment report also contains several other results of international rankings that reveal the programmes' academic reputation.

The committee confirms that the work of educational researchers in RICDE has achieved a strong international reputation during the assessment period, as reflected by levels in international university rankings for research and impact measures for the best publications generated by researchers in both its programmes. The publication records are certainly impressive, with many publications in high- and middle-level journals.

3. Resources and resource policy

Staff numbers at the Research Institute of Child Development and Education (which not only include researchers of UvA1 and UvA2 but also Methods and Statistics staff who contribute to both programmes) have sharply increased over the review period: from 42 persons (equal to 18.78 research fte) in 2006 to 94 persons (equal to 39.68 research fte) in 2011. The number of PhD students (from 10.18 fte in 2006 to 23.44 fte in 2011) and other non-tenured staff (from 2.66 fte to 6.54) has respectively more than doubled and almost tripled over the review period, while the number of tenured staff has grown somewhat more modestly (from 5.94 fte to 9.71 fte). Especially the research capacity of UvA1 has increased over the review period.

The 2009 departure of the Division of Commissioned Research has had a clearly visible effect on the institute's finances. While in 2006 the institute budget amounted to €8,202,000, in 2011 it had shrunk to €7,158,000. The ratio between direct funding and research grants shifted over the review period, especially in the second half (2009-2011). While the share of external money (mostly contract research money, but also some research grants) in the institute's annual budget was in decline (from 50% in 2006 to 35% in 2011), the share of direct university funding rose (from 40% in 2006 to 63% in 2011). These figures and percentages, however, do not take into account that the former SCOKI did not use contract research money to fund fundamental research. Only the – now privatised – Division of Commissioned Research could tap into third stream funding. The Division of Fundamental Research that formed the basis of the current RICDE was, before 2009, funded solely by direct funding and research grants. If only these funds are considered, a threefold increase can be observed between 2006 and 2011, from €2,261,000 in 2006 to €6,826,000 in 2011.

So, even though the departure of the SCOKI Division of Applied Research has brought about temporary budget deficits and particularly reduced the institute's earning capacity for

contract research, the management has been able to guard financial health. Moreover, the exit of this Division of Applied Research allowed the institute to resolve several managerial problems and helped to become more selective by participating only in contract research that contributes to the aims of the research programme. Losing some of its third-stream money may not be problematic, as long as the institute manages to tap into the second money stream, i.e. research grants. In the past, however, success at acquiring NWO and ERC grants has been modest. Enlarging the share of second-stream money is rightfully an important aim of the institute's resource policy. The committee applauds the university strategy of encouraging productivity by providing additional funding resources for staff who obtain external grant income. It also agrees with the strategic decision to divert resources away from the direct funding of PhD students and, instead, redirect funds towards the appointment of post-doctoral academic staff who will be able to obtain research grants, which will also include funding for PhD students. In this way, the overall number of researchers and the available resources for their work will likely increase.

There are notable differences between the programmes in the distribution of funding. While it used to favour UvA2 (Education), UvA1 (Child Development) has consistently been catching up, until 2010. In total, the institute received the exceptional amount of €22,567,300 in indirect funding (research grants and contract research). The former presence of the commissioned research division still resonates within the extremely high earning capacity of €562,776 per tenured research fte per year.

In 2007, the UvA introduced the 'full cost model', in which overhead costs were no longer paid directly but charged to the Research Institutes, and compensated (partly) through matching of research grants. This model has dramatically changed the ratio between personnel costs and other costs (was 82% and 18%, respectively, in 2006, and 58% and 42% in 2011).

The committee established that members of the institute have good access to material facilities for supporting educational and psychological research, such as laboratories, technical support and accommodation. It was pleased to learn that the FMG is the only Social Sciences Faculty in the Netherlands with its own MRI scanner.

4. Productivity strategy

As was mentioned before, RICDE's strategy is to put institutional emphasis on quality rather than quantity of outputs. Even so, the institute's output has increased gradually over the last six years. In total, the research staff of the institute produced 697 academic publications (a yearly average of 4.2 per research fte), 470 of which were published in refereed international journals (with a more or less stable and substantial yearly average of 2.9 per research fte). There was a total of 1065 professional publications, which means an annual average of 6.5 per research fte. Especially the academic productivity varies considerably between the two research programmes, with UvA1 being the more productive one.

In total, the programmes produced 46 PhD theses (28 internal and 18 external). This means that there was an average of 1.2 PhD theses per tenured fte. In absolute terms the number of PhD students was divided more or less equally over the programmes, but when the size of the programmes in fte is taken into account, UvA1 again proves somewhat more productive than UvA2.

5. Societal relevance

To underline the societal relevance of the institute, the self-evaluation report points to the intrinsic social importance of research on family support, child and youth care, interventions in psychopathology (RPCD) and learning disorders, instruction and pedagogy in schools, educational segregation, school dropout, and delinquency (RPEDU).

Evidently, since the exit of the SCOKI Division of Applied Research, a large part of the RPEDU's societally relevant work is conducted in the new Kohnstamm Institute. However, commissioned research that contributes to the aims of the RPEDU research programme continues to be done in collaboration with the Kohnstamm Institute. Besides, RPEDU members are still active in making the results of their educational research suitable for application in various school contexts and available to various groups of professionals (see the programme assessments), and members of the programme hold central societal positions, serve on advisory committees and appear in news media frequently. The same holds for the RPCD, which, given the very positive experiences of RPEDU with the transfer of the commissioned research in the Kohnstamm Institute, is working on the creation of a similar institute for its commissioned research in the domain of child development.

All in all, educational researchers at the UvA show good evidence of the social impact of their research and strong links with policy-makers, practitioners and work through treatment centres such as UvA Minds.

6. Strategy for the future

The institute is very positive about the upscaling of its present programmes. This has not only brought about a better distribution of tasks and a more intensive collaboration among the researchers of the two programmes, it also enables them to anticipate policy changes at all levels, and to respond to increasingly complex and demanding scientific and social problems which demand an interdisciplinary approach. The committee has established that, so far, this strategy has proved successful: the staff, funding and grant-earning capacity of the research programmes have consistently been growing.

The self-evaluation report describes a number of policy measures that together characterise the institute's strategy for the future. First of all, the institute wishes to stimulate its researchers to participate more in the research priority areas (RPAs: RPA Brain and Cognition; RPA Affect Regulation; RPA Institutions and Inequality; and RPA Urban Studies) that have been defined by the UvA and the FMG – thereby qualifying for a larger share of RPA-associated funding. As of yet, the institute has barely profited from this type of funding, mainly because it was not a serious point of attention in the period before the reorganisation, during which several professors retired and the institute was led by ad interim directors. This situation is expected to change shortly. During the site visit, the institute management announced the intended creation of a new interdisciplinary RPA, whose topic relates well to RICDE research. Just after the site visit this RPA, entitled Yield, was indeed established. Yield will cover research on human development, from infancy to young adulthood, with perspectives from Medicine, Psychology, Pedagogics, Education, Communication, Economics and Psychometrics. The faculty funding that is available to start the new RPA (2.4 million euro) will be used to appoint twelve young researchers to advance the institute's research in this particular area and to assist in acquiring additional funding to extend the RPA.

Although the institute has obtained substantial indirect funding during the assessment period, it has not acquired many high-profile individual grants (such as NWO Vidi and Vici grants or ERC starting and advanced grants). The institute hopes that by obtaining such grants, it will

not only further increase its funding, but also its profile both within and outside the Faculty and university. To assist individual researchers and groups with the preparation of grant proposals, the university and faculty have established offices that provide financial, juridical and technical advice and support.

Furthermore, the institute encourages collaboration with societal partners in academic and non-academic health care, in academic training schools, in commissioned research, and in government. These partnerships are also thought to bring about new funding opportunities for applied as well as for fundamental research.

Finally, to further stimulate the preparation of its students and to enrich its scientific output, the institute will intensify the collaboration with the graduate school, so that (research) master students and PhD students can participate in the research programmes more intensively.

In conclusion, the committee notes that several factors (the fact that the government is reducing direct funding of research, that universities and faculties invest their money primarily in RPAs, and the departure of the Kohnstamm Institute for applied research) have resulted in a situation in which programmes and researchers have become almost entirely dependent on insecure indirect funding. Combined with an increase in their teaching load, this puts high pressure on them. According to the committee, the self-evaluation report rightfully mentions this as a major threat for the future (even though the appointment of a substantial number of new (postdoc) researchers in the context of the recently newly established RPA Yield may help to alleviate to some extent the teaching load of the senior staff).

7. PhD training and supervision

According to the self-evaluation report, the institute and the Graduate School of Child Development and Education (GSCDE) closely collaborate to offer PhD candidates a training programme that prepares them for a successful research career inside or outside academia. This programme consists of (1) courses that offer advanced study of the subject of the PhD study, (2) courses in research methods and statistics, (3) other supporting courses (e.g., an academic writing course). In consultation with their supervisors, PhD candidates decide on their individual learning trajectories. They can choose courses from the (research) master education programme, from other graduate schools within the faculty, and from courses offered by the National Research Schools (NRS; e.g. ISED, ICO, IOPS). The Graduate Studies Committee must approve the individual education programme.

According to the PhD students that the committee spoke to, one of the national doctoral schools (ISED) has been rather inactive for some time, but this problem has now been solved. The PhD students from RICDE highly appreciate the interaction with their peers from other institutes and universities in the context of these National Research Schools. They are also pleased with the local Graduate School and the courses it offers, although some would prefer more attention being paid to training in didactical skills and in writing grant proposals. The PhD students mentioned that interaction between them is stimulated in various ways, e.g. through seminars, colloquia and symposia that are organised at different levels. They generally like the interdisciplinary nature of many of these meetings and the stimuli to interact and cooperate with students from different programmes. Some students indicated that they would welcome more links between research and relevant practice outside the university sector to enable career opportunities in the wider world.

PhD students are allowed to visit one international research congress each year, on condition that they present their ongoing research. Fees, travel, accommodation, and expenses are all

paid for by the research institute, if not already covered by the specific research project budget. Special additional requests can be honoured by making an appeal to the Stichting Kohnstamm Fonds for dissemination of educational research.

PhD students are supervised by at least two experienced researchers (associate or full professors), and most of them have an (additional) daily supervisor. The PhD student fully participates in the research activities and lab meetings of the supervisors' research lines, and relations with academic staff do not seem inhibited by hierarchy. Typically, PhD students meet their daily supervisor very regularly (once a week), while the meetings with the senior supervisor(s) are typically scheduled every three or four weeks. Most students are pleased with the availability of their supervisors and the guidance they get from them, although some comment that they would have preferred more frequent contact. Every year, there is a formal performance interview with the programme leader, in which the PhD student and supervisors discuss the progress report (written by the candidate), the progress of the research project, and the supervision and education of the candidate.

Out of the 26 funded PhD students who enrolled between 2003 and 2007, 35% had graduated after four years. After five years, another 15% had graduated and after six years another 12%. While a total of 62% of PhD students had finished their projects after six years, 30% was still working on them, and 8% had dropped out. According to the institute management, the relatively long completion time can be explained by pregnancies, part-time employment (0.8 appointment for 5 years) and also by the fact that finishing within the period of the contract was – until recently – not considered a hard target, as extensions were frequently granted. However, because of financial constraints the PhD completion culture has changed, and there is now strong pressure to complete the thesis in time.

Except for PhD completion times, which has recently already become a serious point of attention, the committee thinks that the PhD training and supervision offered by the institute are very well organised and of very good quality.

B. Programme level

The committee comes to the following overall programme scores for quality (Q), productivity (P), relevance (R), and vitality/feasibility (V) for the two programmes of The Research Institute of Child Development and Education of the University of Amsterdam:

Code	Programme name	Q	P	R	V
UvA1	Child Development (RPCD)	4.5	5	4.5	4
UvA2	Education (RPEDU)	4.5	4.5	4	4

The detailed assessment for each programme follows in the next section of this report.

Programme UvA1: Child Development (RPCD)

Programme director: Prof. Geert-Jan Stams

Research staff 2011: 3.6 tenured fte, 16.5 total fte

Assessments:	Quality:	4.5
	Productivity:	5
	Relevance:	4.5
	Viability:	4

Brief description

The programme focuses on normal and abnormal development, as well as interventions that support healthy child development. A central focus that unites the programmes of research concerns parenting and caregiving for children. Programme goals include understanding normal and abnormal development, and developing interventions that address critical issues related to parent-child interaction and child psychopathology. More specific topics include: traditional and non-traditional families; non-parental institutional care; the role of the father; the association between temperament and parenting; and interventions addressing child-rearing problems, psychopathology and juvenile delinquency.

This programme was formed in 2009, when a larger reorganisation occurred at the institute level as a number of professors left the university and new professors were hired. In addition to being part of the institute, the programme is also embedded within the RPA Brain and Cognition and RPA Affect Regulation.

Quality

The quality of the programme is excellent. It includes very strong researchers who make key contributions at an international level. The areas covered by the programme are rich and diverse, and include both basic research and research focused on the clinical needs of children and families with difficulties. The research staff has collaborations with international colleagues. Key publications are in good to excellent international journals.

The broad and vast range of topics included in the programme creates both opportunities and challenges. Although the various areas included could create possibilities for synergy across diverse areas, greater programme and institute cohesion may enhance these opportunities.

As documented in the institute report, the acquisition of external grants has been relatively successful.

Productivity

Since 2006, the total research staff of this programme has more than doubled, from 6.22 fte in 2006 to 16.45 fte in 2011. This particularly reflects an increase in the number of PhD students. In terms of productivity, this programme is exceptional. The number of publications is very high, as is the number of citations. The number of academic publications was 385 (an annual average of 5.7 per research fte), 265 of which are English refereed articles (4.0 per research fte). The number of professional publications is very high (7.5 per research fte).

The number of PhD theses is also very high: over the review period 24 projects (15 internal and 9 external) were completed. This amounts to an annual average of 1.44 per tenured fte.

Societal relevance

The programme is excellent with regard to societal relevance. Its staff has developed and evaluated practice protocols that are being implemented at the national and international level. These include a cognitive behavioural treatment protocol for anxiety, a mindfulness approach for children with externalising problems, an attachment-oriented intervention for foster parents, and assessment instruments for use in (forensic) child and youth care. A clinical setting directly associated with the institute ('UvA Minds', directed by Prof. Bögels) provides a very rich setting in which clinical practice with families and children is combined with conducting high-quality research. The faculty also has a presence in the media.

Viability and feasibility

There have been many changes in the programme in the years since the previous evaluation. Following an institute reorganisation in 2009, it quickly established itself and appears to be well supported at the institute and university levels. The programme has developed significant connections to the university level RPA initiatives on Brain and Cognition and on Affect Regulation. The committee believes that the establishment of RPA Yield is a positive development in this regard. On the other hand, the programme could profit from a more unified model of leadership that would provide greater integration and cohesion across the diverse areas of research. Greater integration/coordination across research projects and ideas is likely to maximise study quality and impact. The current need for greater cohesion, a disproportional reliance on third-stream funding, and an anticipated decrease in availability of external funding represent challenges for the future.

Conclusion

The programme, established in 2009 in response to changes in staff and the desire to consolidate, has had a successful launch. The quality of its research and the productivity are excellent, as is its societal relevance. The programme's faculty members make key contributions at the national and international levels. Paying more attention to the integration within this programme and of this programme with others in the institute, as in the recently established RPA Yield, may strengthen it further.

Programme UvA2: Education (RPEDU)

Programme director: Prof. Peter Sleegers (until 2010); Prof. Monique Volman (from 2010)

Research staff 2011: 5.6 tenured fte, 21.0 total fte

Assessments:	Quality:	4.5
	Productivity:	4.5
	Relevance:	4
	Viability:	4

Brief description

This programme is primarily concerned with understanding how the educational process contributes to the development of cognitive and emotional skills. The research carried out is distinctive in addressing factors at the micro, meso and macro levels in primary and secondary education. It aims to provide theoretical foundations for the improvement of teaching and learning and to support educational innovation. There are two main lines of research, concerning learning outcomes and learning processes.

Quality

This programme has made distinctive contributions to the testing of educational theory, covering diverse topics such as literacy development, the scaffolding of learning, citizenship competences and student motivation, using a range of methodologies. Members of the programme are internationally visible and prominent, serving on editorial boards and boards of international organisations and conferences, and a substantial proportion of their publications has appeared in high-impact forums. They have been successful in acquiring research grants from competitive and prestigious sources such as NWO and VO Raad (Educational Council).

Productivity

In terms of bibliometric outcomes, the programme scores are good, though some areas of research are linked to journals with relatively low impact factors. Over the assessment period as a whole, the quantity of research outputs has been very good (312 academic publications), with a fair rate of productivity (3.5 annual publications per research fte), and with 205 publications in international refereed journals (2.3 per research fte). During the assessment period, the number of academic publications has increased substantially, in most categories. The number of professional publications is very high (6.1 per research fte).

Numbers of PhD students have doubled throughout the assessment period. The number of PhD theses completed in the assessment period is high at 22 (13 internal and 9 external; 0.94 per fte for tenured staff).

Societal relevance

The research of the programme has made an impact on national educational policy and practice. It has been funded by several government organisations and has generated tests that are widely used in schools. The dissemination of research to wider audiences is well represented by professional publications, handbooks and high-profile presentations by staff at relevant events. Members of the programme chair or participate in policy-making committees, and the programme research has generated some media attention. There is regular contact between programme researchers, educational practitioners, teacher trainers and the general public.

Viability, feasibility and vision for the future

The programme's viability is very good, with three new chair appointments having been made and a steady stream of funding for new projects and good PhD student recruitment. The research plans for the future seem very coherent, being well focused and with a clear direction. The programme has developed significant connections to four RPAs, and RPEDU is now (also) part of the newly established RPA Yield.

On the other hand, the programme is confronted with some threats, such as the relatively low number of tenured staff in proportion to non-tenured and PhD staff members, which is likely to cause work overload for the former and could limit the impact of the research. The recent establishment of the new RPA Yield has led to a substantial number of new appointments of (post doc) researchers, with more to come, which should alleviate to some extent the teaching load of the senior staff.

Conclusion

This is a very successful programme which is recognised nationally and internationally for making a strong contribution to educational research, and which impacts significantly on policy and practice in the Netherlands. Its work has coherence and clear direction for the future. The recruitment of more tenured staff, improved completion rates for doctoral students and increased levels of external funding would help to ensure its viability.

8. VU UNIVERSITY AMSTERDAM

A. Institute level

1. The institute

The Educational and Pedagogical Sciences programmes at VU University Amsterdam are part of the Faculty of Psychology and Education (FPP). Their core mission is ‘to generate insights into practices and behaviours in education, parenting, and professional care and their linkages to individual differences in child development and functioning.’ This mission is pursued by research in and together with the field of regular and special education, youth care (mental health), and care for children with illnesses and disabilities.

The research is subdivided into three programmes:

- Theory and research in education (VU1, Department of Research and Theory in Education, since 2006);
- Challenges to child-rearing relationships (VU2, Department of Clinical Child and Family Studies, since 2006);
- Educational Neuroscience (VU3, Department of Educational Neuroscience, since 2009).

The most important changes since the previous national Research Review of Educational and Pedagogical Sciences (2006) are (1) the installation of the research programme on Educational Neuroscience and (2) the merging of the programmes on Developmental Education in the School Context and Education, Morality, and Religion into the new Theory and research in education programme (following the suggestions made in the 2006-2008 Midterm Review).

In accordance with university policy, the Educational and Pedagogical Sciences programmes participate in a number of university-wide multidisciplinary institutes, such as EMGO+ (on health and care research), Visor (on philosophy and theology), and Phoolan Devi (on crime and punishment). The Departments themselves have taken the lead in LEARN! (on education, established in 2009).

The committee was pleased to learn that education is one of the university-wide profiling themes of the VU University. This, as well as the active participation, and in the case of LEARN! prominent, in the above-mentioned multidisciplinary institutes is seen as both a strength and an opportunity.

2. Policy on scientific quality and academic reputation

The committee has established that the institute has a well-articulated strategy for achieving excellence. During the assessment period, the institute has set strict and ambitious criteria for appointments and promotions. Individual staff members are expected to publish frequently in high-impact journals, acquire grants, act as supervisors of many PhD theses, etc. Well-defined, specific criteria are provided for each step of their academic careers. It is also faculty policy to assess, at the start of and during projects, the quality of the proposed methodology, the research’s relevance, and its compliance to ethical, verifiable, and reliable research standards. More generally, ethical issues take a remarkably important place in the institute’s research policy.

The committee has seen evidence of very strong leadership within the institute. There is an impressive level of central coordination between the three individual programmes, while, at the same time, programme leaders are given quite a lot of autonomy. As was explained during

the site visit, the chairs have delegated responsibility for research time and capacity, but their decisions have to be in line with the Faculty Board's resource policies. Formal quarterly meetings between programme leaders enhance the synergy between the programmes. The smooth way in which the newly developed Educational Neuroscience programme has been incorporated into the institute can be taken as a clear indication of this synergy.

Indicators of the academic esteem of senior staff members listed in the self-evaluation report include editorial positions at respected national and international journals, memberships of national and international grant and awards committees, as well as scientific governance bodies, and regular invitations to deliver keynote lectures at international conferences.

The recent citation analysis by Prins shows that the publications VU1 and VU2 are cited either at or slightly above the average rate of citation for the respective subfields, while the VU3 programme scores far beyond its reference points.

3. Resources and resource policy

Staff numbers at the Educational and Pedagogical Sciences programmes have more or less doubled over the review period: from 36 persons (equal to 16.3 research fte) in 2006 to 71 persons (equal to 32.8 research fte) in 2011. This was mostly due to the start of the new and successful Educational Neuroscience programme in 2009, which currently involves, besides the two programme leaders (both coming from Maastricht University), mainly PhD students (10.2 of the total of 12.7 fte in 2011). The total of tenured staff has not substantially grown with the arrival of the third programme.

The increase in research staff members was paralleled by a substantial increase in total funding, from €2,236,892 in 2006 to €4,195,246 in 2011. The ratio between direct funding, research grants and contract research has shifted over the review period, especially during the last two years (2010 and 2011). While in 2006, 73% of the total budget was obtained from the university itself, in 2011 direct funding amounted to only 44% of the total income. In the last year of the review period, research grants and contract research (a category virtually absent in 2006) brought in 40% and 16% of the annual budget, respectively. There are remarkable differences between the three programmes in this distribution, with VU2 and particularly VU3 being responsible for the majority of the incoming indirect funding. In total, the three programmes received €6,138,947 in indirect funding (research grants and contract research) over the review period. This makes for an earning capacity of €168,652 per tenured research fte per year.

The institute considers grant acquisition as important and even critical to the viability of research programmes and staff careers. Sources during the review period included NWO LEARNING, NWO-MaGW project and programme grants, ZonMW and programme grants, European Union grants (FP6), and grants from charitable foundations. The 2011 VICI grant to one of the programme leaders of VU3 further underlines the already impressive earning power of the Educational Neuroscience programme. Although the self-evaluation report identifies the general economic slowdown and diminishing funds from university and national funding agencies as threats, it also points out funding opportunities. By being part of the EMGO+ institute, Educational and Pedagogical Sciences has access to health research funding, which was elected one of the ten focus fields in the Netherlands and is an important area for EU research funding. Educational neuroscience is a possibly attractive domain for EU research funding, too.

By allocating direct funding based on the earned research grants as well as contract research, the institute opts for rewarding the programmes' general earning capacity. While the weighing of contract research attests to the importance it attaches to addressing practice- and policy-relevant research, a threat to general quality may be that contract research can vary in quality and quality control.

4. Productivity strategy

Over the last six years, the research staff members of the three programmes produced 817 academic publications (a yearly average of 6.6 per research fte), 324 of which were published in international refereed journals (which amounts to a high yearly average of 2.6 per research fte). Whereas the absolute academic output of the joint programmes has increased over the review period (from 117 publications in 2006 to 140 publications in 2011), the input-output ratio was somewhat less favourable at the end of the review period. When academic production is examined at the programme level, it becomes clear that there are great differences in terms of overall productivity, but also in terms of the nature of the preferred academic publications. In addition to academic publications, the institute produced a total of 444 professional publications, amounting to a high annual average of 3.6 per research fte.

In total, over the review period, the programmes produced 31 PhD theses, which amounts to 0.9 PhDs per tenured research fte. Again, taking into account the available fte for tenured staff, the PhD-generating capacity of VU3 is the institute's highest.

5. Societal relevance

The self-evaluation report notes that the Faculty Board strongly encourages its departments to embed their research in long-term commitments with societal stakeholders in order to raise the level of education, caregiving and participation, especially for challenged populations. The faculty supports this aim by equitable allocation of its direct funds on the basis of third-stream money. To the best of the committee's knowledge, the FPP of the VU is the only faculty in this assessment round that applies such a direct funding policy.

The committee has established that expectations for staff members are set high where societal relevance is concerned, leading to a high output of professional publications and extensive relations with the field. The Educational and Pedagogical Sciences Research programmes foster over 20 long-term partnerships with societal organisations in the field of care and education. Sharing of knowledge takes place at presentations and workshops given in the field, through shared appointments, academic workplaces, and funding by more applied competitive programmes like ZonMW or NWO PROO, as well as by direct investments from the field. Every year, interviews are given to major national newspapers, magazines, radio and television. For more details, the committee refers to the specific programme assessments.

Several programme members have been rewarded for their societal investments. In 2011, the director of VU3 (Educational Neuroscience) received the VU University Societal Relevance Award. VU2 (Challenges to childrearing relationships) has also seen its members nationally rewarded for their contributions to the betterment of lives of children with disabilities and child maltreatment victims.

6. Strategy for the future

The faculty has developed a three-pronged strategy for the future. Firstly, FPP has proactively embraced VU's policy of multidisciplinary research institutes by actively investing in the participation in the above-mentioned interdisciplinary institutes EMGO+ and LEARN!. By

doing so, the programmes share in the university's strategic funding, have direct access to the best infrastructure and research groups, and are part of VU's strategic planning. However, the governance of these multidisciplinary institutes puts extra pressure on the precious time of the senior researchers. The 2008-2011 Midterm Review encouraged further expansion of the LEARN! Institute, which now suffers from limited participation from other faculties. As a result, talks are underway with new groups from the Faculties of Social Sciences, Medicine, Humanities, Natural Sciences, and Philosophy.

Secondly, in its budget allocation to departments, the FPP aims to clearly and directly reward productivity and earning capacity. Departments receive yearly lump sums based on the number of dissertations produced, scientific personnel on external research funds, and course credits. They are given the freedom to spend these budgets. Using tight prognosis and planning, this allows departments considerable strategic room to build and renew their groups, and prevent reductions in research capacity.

Thirdly, FPP has an ambitious tenure track system in place that provides a strong impetus for aspiring staff to build their academic reputation. As a result, increasing numbers of senior and junior staff have been active and successful in acquisition, which provides a strong basis for the viability of the research programmes.

The committee has been informed that there are no plans for further aggregation of groups. VU1 (Theory and Research in Education) might, however, grow in 2012 and 2013 by incorporating research from the university's teaching education department (CETAR). A decision on the potential merger of the research groups had at the time of the site visit not been made. According to the committee, it remains to be seen if a merger would contribute to the dynamism and coherence of the newly established programme VU1, which is still in the phase of bringing together two quite different lines of research in terms of research themes as well as theoretical and methodological approaches.

The senior staff base for the Educational and Pedagogical sciences was increased through new chairs in 2012 (around child abuse and interpersonal relationships) and in 2013 (around issues in special education and neuropsychology). The latter (upcoming?) appointment was made possible because VU chose the director of VU3 (Educational Neuroscience) as one of the three University Chairs for high-profile scholarship. This can be considered a strong sign for the appreciation and respect that the Educational and Pedagogical Sciences in general and VU3 in particular can count upon within VU.

7. PhD training and supervision

At the faculty level, all research training of PhD students is organised within the William James Graduate School (WJGS). PhD students are offered a tailor-made package of courses, 40% of which focus on advanced data analysis and transferable skills (e.g. ethics, writing & presentation, career development, time management) taught at the local level, in on-campus centres or by WJGS research staff.

The remaining 60% of the courses is conducted at various KNAW-accredited National Research Schools (NRS), such as ISED. The Graduate School's director reviews the training and education part and makes sure that the WJGS curriculum committee has reviewed the proposed courses. The WJGS examination committee ensures the proper completion of the course work in the PhD programme. The PhD students that the committee spoke to were very pleased with the current arrangement, whereby part of their PhD training takes place in the local graduate school and the other part in the national doctoral school.

The faculty has set aside a dedicated budget of €150,000 per year for its so-called Talent Programme, which is aimed at the education and stimulation of the PhD students. This budget allows PhD students to take extended visits at the many international collaborators of the FPP research staff. To increase the percentage of PhD students who graduate on time, the Talent Programme awards a bonus when a PhD student finishes on time. Finally, the Talent Programme rewards departments that successfully guide a PhD graduate to a VENI or comparable personal fellowship.

For each PhD project, a training and supervision plan has to be submitted. The scientific and ethical review committee subsequently checks this for scientific content and feasibility. PhD students have ample opportunity to meet with their supervisors. Except for PhD students of the VU1 programme, most others have more than one supervisor, including a daily supervisor.

During the site visit, PhD students indicated that they are very pleased with the scientific and social climate at the institute. There is, in their view, a good mixture of various types of scientific meetings (i.e., monthly meetings with all staff members at the programme level, monthly meetings with all the PhD students, frequent meetings with the supervisors). Moreover, there is an active PhD social committee that helps PhD students to find their way and feel at home in the institute and university.

Out of the 35 PhD students who enrolled between 2003 and 2007, 37% had graduated after four years. After five years another 14% had graduated, and after six years another 9%. After 7 years, 29% had not yet finished their projects, while 9% had dropped out. It is to be expected that the above measures in the context of the Talent Programme will lead to an improvement of these graduation rate data.

B. Programme level

The committee comes to the following overall programme scores for quality (Q), productivity (P), relevance (R), and vitality/feasibility (V) for the three programmes at the Faculty of Psychology and Education of VU University Amsterdam:

Code	Programme name	Q	P	R	V
VU1	Theory and Research in Education	4	4	4	3.5
VU2	Challenges to child-rearing relationships	4.5	4	5	4.5
VU3	Educational neuroscience	4.5	5	4	4

The detailed assessment for each programme follows in the next section of this report.

Programme VU1: Theory and Research in Education

Programme directors: Prof. Doret de Ruyter and Prof. Bert van Oers

Research staff 2011: 2.3 fte tenured, 15.0 total fte

Assessments:	Quality:	4
	Productivity:	4
	Relevance:	4
	Viability:	3.5

Brief description

This programme focuses on identifying the best educational practices for developing children's capabilities for integrating with and contributing to modern society (as expressed in the subtitle of the programme: 'Education for the good life in a diverse society'). Its research involves two main themes:

- Ideals and identity;
- Meaningful education and citizenship.

The former incorporates a particular interest in children's moral development and the influence of religious aspects of culture on this, and the second incorporates a particular interest in the roles of play and inquiry in children's learning from a developmental education perspective. Overall, the programme is characterised by a concern to integrate theoretical and empirical strands of research.

Since 2006, the total research staff has decreased in fte terms, due to the loss of tenured staff and doctoral students.

Quality

Following the recommendations of the last assessment, this programme was created in 2009 through the merger of two previous programmes. During the assessment period, strong contributions have been made to research related to both the themes listed above, and particularly in relation to the first (for example, through the REDCo project funded by the EU). The quality of this research is reflected by the quality and diversity of the submitted publications, which include edited collections and articles in high-quality, international refereed journals. Senior staff members have served on the boards of leading journals in relevant fields and research associations, given invited keynotes in international conferences and taken up visiting professorships abroad.

The programme leaders have strong international reputations. Grant success, especially contract research, is very good. Members of the programme have been successful in acquiring research grants from competitive and prestigious sources such as the EU KP6 programme and NWO, with the useful and applied value of their research being recognised by government grants such as from the Ministry of Internal Affairs. However, the overall amount of indirect funds acquired over the assessment period (both in absolute and relative terms) is rather modest.

The two former programmes are presented as a joint new programme under a common general theme. Although there are already some visible connections between the two thematic strands, available evidence suggests that their full integration is a task for the future. Standard

impact scores are not very high, but this may be related to the strong theoretical orientation of the programme.

Productivity

In terms of bibliometric outcomes, the programme scores are very good, given the small number of tenured staff: 471 academic publications (9.79 per research fte) were produced, but only a relatively small proportion of them appeared in international refereed journals (96; 1.99 per research fte), while many others were published in other academic publications (frequently national). The number of professional publications is also very high (163, i.e. 3.38 per research fte).

The number of PhD students has dropped slightly through the assessment period and is now quite small (from 4.8 fte in 2006 to 2.1 in 2011). The overall number of PhD theses completed in the assessment period is good (12; 0.72 per fte for tenured staff), but this number has also dropped in recent years.

Societal relevance

While the programme obviously aims to enhance academic knowledge, its members invest significantly in the valorisation of their research too. The programme's research is very effectively disseminated to social policy-makers and practitioners nationally, and quite effectively internationally also. This is not only true for those investigating Developmental Education (which has strong ties to the Dutch schools, school advisory institutes, and teacher training institutes that have implemented this type of education), but also for the philosophical and historical researchers who work on societally relevant topics (such as multi-ethnic diversity worldview education). Programme members give a high number of lectures for lay audiences, and their work features regularly in newspapers and other media. The appointment in 2009 of a well-known researcher in education in migrant families, with a particularly strong track record in valorisation of research, has further contributed to the societal relevance of the programme.

Viability, feasibility and vision for the future

The programme's distinctive contribution to Dutch educational research, the strong reputation of its senior members, and their recognised role in shaping national educational policy and practice should help to ensure its continued viability. New opportunities for interdisciplinary research, for example with neuropsychologists, look promising. On the other hand, the programme is confronted with some threats, such as the relatively low number of tenured and senior staff and (internal) PhD students, the rather unfavourable age composition of the staff (although the number of young staff members on tenure tracks has already started to increase in 2011 and is anticipated to further increase afterwards), and the limited level of achievements in research grant acquisition. The merger of the Department of Theory and Research in Education with part of CETAR may provide opportunities for new lines of research. It may, however, also create some uncertainty about the organisational strength and programmatic coherence. Exploring in more detail if, how and when the CETAR research programme can be productively merged with the current programme, is therefore an important task for the near future. But if the current shortage of younger staff and (internal) doctoral students can be further overcome, grant levels can be maintained, and the merger with CETAR can be successfully done, the resources for effective future development would seem to be available.

Conclusion

This relatively new programme is recognised nationally and internationally for making a distinctive contribution to educational research (theory) in both thematic strands of its activity. As well as the above-mentioned threats that need to be addressed, the growing connections between the two thematic strands need to become more visible.

Programme VU2: Challenges to child-rearing relationships

Programme director: Prof. Carlo Schuengel (since 2006) and Prof. Catrin Finkenauer (since 2012)

Research staff 2011: 2.5 tenured fte, 13.2 total fte

Assessments:	Quality:	4.5
	Productivity:	4
	Relevance:	5
	Viability:	4.5

Brief description

This programme is aimed at studying and contributing to the quality of children's relationships with caregivers and peers in a variety of challenging situations. Multiple foci can be distinguished, including the transition to parenthood, challenged parenthood, children's special caregiving needs, and interpersonal antecedents and consequences of victimization. Research is inspired by different theories and frameworks related to attachment, interpersonal relationships, neurobiology and genetics, disabilities, psychotherapy, and child welfare. The programme is embedded in several multidisciplinary research institutes at the university level (EMGO+ on health and care research, LEARN! on education).

Quality

The programme leaders have a clear vision of the programme's mission and its scientific and societal relevance. They show strong leadership within their team and are committed to advancing science through conducting methodologically sound and theory-based research in diverse and challenging populations. The programme successfully integrates diverse theories and backgrounds (e.g., attachment theory, bio-psychological and behavioural approaches) to shed light on issues of great importance to the field of (ortho)pedagogics (e.g., concerning disability, foster parenthood, victimisation). Research findings are published in high-impact journals, such as the *Journal of Consulting and Clinical Psychology*, *Journal of Child Psychology and Psychiatry*, and *Journal of the American Academy of Child & Adolescent Psychiatry*. Programme leaders collaborate with excellent research groups nationally and internationally. Their international reputation is also evident from the invited lectures, associate editorships, and awards. The programme has competed successfully for research grants within open and more focused programmes of NWO and ZonMW, and also manages to secure funding from charity funds and field organisations. The number of PhD students in particular has grown tremendously in the past years.

During the evaluation period, total funding has increased from €1,290,000 to €1,620,000; this is due to the increase in research grants and contract research. The direct funding has dropped during the evaluation period, proportionally, but in the final year also in absolute terms.

Productivity

Overall, the programme has a very good academic research output, given its size and the number of tenured research staff in particular, with a total of 186 academic publications (an annual average of 3.89 per research fte and 12.24 per *tenured* staff research fte). The number of English peer-reviewed articles (85 in total, an annual average of 1.78 per research fte or 5.59 per *tenured* staff research fte) has grown substantially over the review period, although it varies from year to year (related to the small size of the programme). The number of professional publications is very high: 186 or an annual average of 3.89 per research fte.

The total number of completed PhDs is 8 (0.52 per tenured fte). This figure is comparatively low, but should be viewed in light of the fact that the programme had only one full professor during the review period.

Societal relevance

Throughout the self-evaluation report and the research assessment in general, societal relevance was demonstrated to be one of this programme's central ambitions and strengths. The programme's mission involves the support of children and families at risk as a central aim, as well as the broadening and enrichment of research and theoretical perspectives to include these challenging populations and situations. As a result, the programme is embedded strongly in the field of health and care providers. These structural collaborations have provided key resources in terms of finances and personnel, and thus support the programme's research and clinical capacities. Substantial grants have been obtained to conduct combined research- and practice-oriented research (e.g., in the field of child abuse). The programmatic pursuit of a research-based approach of relational problems and opportunities in the fields of disabilities, child welfare, and psychotherapy and the numerous studies investigating the effects of clinical interventions and assessment methods have influenced and improved practice. Because of this explicit and consistent focus, the programme's societal relevance is judged to be excellent.

Viability, feasibility and vision for the future

This programme is rather small in terms of tenured staff. During the evaluation period there was one full professor (Schuengel); since 2012, a second full professor has been appointed (Finkenauer). The number of PhD students increased from 4 (2.6 fte) to 13 (7 fte), as did the number of non-tenured staff (from 1.4 fte to 3.7 fte, although it varies from year to year).

Recently, five new assistant and associate professors have been appointed, including one tenure-track candidate. In the committee's opinion, the selection of excellent new staff members with complementary expertise, the strong shared vision and leadership, and the collaborative climate contribute to the programme's strong viability.

Productivity is expected to increase given the new appointments and the recent strong increase in PhD students. Also, the gradual building of data sets (longitudinal) and the interdisciplinary collaborations established (e.g., within EMGO+) suggest good prospects for the programme's future research and output. The appointment of a second chair of the programme seems a very good decision, given the heavy load of multiple tasks and the high number of PhD students in the programme. If this proves insufficient, additional appointment of tenured staff may be considered. Also, the decreasing and fluctuating direct funding should be carefully followed up and corrected when necessary to ensure the programme's vitality and future.

Conclusion

The programme is judged as excellent in terms of its societal impact and relevance, vision, leadership, and collaborative spirit. It should be commended for its continued use of time-intensive methods and populations that are difficult to reach, and for its theory-based approach to professional problems. Despite the challenges, the programme has increased its output and funding. Although the number of PhDs has been limited, the steep increase in PhD students in recent years and the appointment of an additional full professor offer good prospects for future productivity.

Programme VU3: Educational Neuroscience

Programme director: Prof. Jelle Jolles, Prof. Lydia Krabbendam (since 2009)

Research staff 2011: 1.5 tenured fte, 12.7 total fte

Assessments:	Quality:	4.5
	Productivity:	5
	Relevance:	4
	Viability:	4

Brief description

This programme is new, begun in 2009 when the programme leaders were recruited to VU from the University of Maastricht (Prof. Jolles arrived in 2009; Lydia Krabbendam arrived in 2009 and was promoted to professor in 2011). Its mission is to further fundamental knowledge of learning and to translate this knowledge into educational practice via a multidisciplinary approach. The main objectives can be summarised as (1) deepening understanding of normal, exceptional and subnormal learning, (2) translation of insights from neuroscience areas relevant to education, (3) bridging studies in the neuroscience-education interfaces, and 4) contributing to the development of new educational programmes, including the neuroscience-related perspectives.

Quality

The programme is still quite new, and thus few of its many publications to date (through 2011) have been focused on the interface of education and neuroscience. However, its leaders have an impressive history of very high-impact publishing tradition in the more mono-disciplinary history of their work in basic neuroscience. Their recent publications have moved increasingly to the education-relevant domains while maintaining the goal of publishing in high-impact forums. From this perspective, the programme seems to possess very high quality and great potential. Also, the training of PhD students emphasises quality. The claim that the PhD student ‘typically submits at least five papers in refereed journals’ reflects the ambition of the research training, which, to be successful, requires careful participation from the few senior faculty members.

Active connections to other education-related entities of the university are emphasised. The senior leader, Prof. Jolles, directs a NWO-funded LEARNING programme of 40 NWO-funded researchers. Collaboration with a number of research units and schools, as well as education-related industries and publishers, is of central interest. Leaders have been active in international actions such as OECD-CERI Brain, Learning and Education and one of its ‘products’, Education of the International Mind Brain and Education Society, which has helped them to become ‘well-networked’ to relevant international collaborators. There are strong international connections to high-quality neuroscience labs in UK, China and US.

Productivity

The quantitative documentation describes the results of a programme that started in 2009 and that consists so far of a small staff in combination with an increasingly large number of PhD students. Nevertheless, the programme has already attracted a very high level of external funding. The same ambition and success are reflected by the output, which amounts to a total of 160 publications (an annual average of 5.94 per research fte, but of 34.78 per tenured research fte). The number of English peer-reviewed articles is particularly impressive (143 in total, an annual average of 5.31 per research fte, and of 31.08 per tenured research fte). The

number of professional publications is also good: 95 or an annual average of 3.53 per research fte.

What is remarkable to the committee is that a substantial portion of the research in the programme is approaching the more typical activity of behavioural sciences – large-scale studies where productivity may demand substantially more time and effort.

Altogether, 11 PhDs (all external) supervised by the senior faculty of the programme have defended their theses during the review period, which amounts to an impressive annual average of 2.39 PhD completions per tenured research fte. Notably, all of these students had started their PhD research elsewhere. Many more completions are scheduled for the near future.

Societal relevance

Social relevance is identified as a core value in the programme's mission. This emphasis has motivated the research staff to contribute to R&D leading to educational interventions for public use. A most remarkable confirmation of this is that Prof. Jolles recently received the VU's Societal Impact Award. The visibility of the senior faculty in public forums and media associated with education is high. The programme hosts the Centre for Brain and Learning, which organises workshops on brain and learning for teachers and parents. Because the programme is so young, its most notable achievements are works in progress. In any case, it has shown a very good start.

Viability, feasibility and vision for the future

At the time of assessment, the programme had existed for less than four years. However, during this period much has already been accomplished in demonstrating the viability of the approach. Thus, the programme has had a very good beginning. Its productivity and funding basis seem to be almost without comparison. It consists of a very good and highly motivated group of senior experts, albeit relatively small, and is able to recruit a lot of PhD students. Moreover, the programme already acts as an 'attraction' for many other researchers, both within and outside the institute. The very strong background expertise, high motivation, and optimistic spirit of the senior faculty concerning the opportunities of educational neuroscience place them in a situation in which they have the potential to make a remarkable contribution to the production of innovative ideas and applications of neuro-scientific models to education. This spirit seems to have infiltrated the younger generation of PhD students, which ensures successful building of new kinds of multidisciplinary collaboration. It is likely that the beneficial effects will not be restricted to the programme as such, but will stimulate representatives of different domains of educational sciences – both within and outside the institute – in many senses.

Despite the generally enthusiastic report above, there are also some concerns. The committee is not fully confident that the leaders have sufficiently thought through their opportunities to take further steps. In spite of their impressive output, external funding, and networking activities, and of the presence of some references to their strategic approach in their self-assessment report, they were not able to articulate a clear and convincing strategic plan or overall approach. The answers to the committee's affiliated questions regarding concrete projects that would directly benefit the curriculum, students, and teachers also did not reveal concrete long-term planning. It appears that the high supervision load and short history of contemplation of the possibilities have so far somehow inhibited the development of a clear and convincing strategic research plan.

Conclusion

This is a new and highly promising programme. Both its quality and productivity are already exceptionally good. The promising start, however, still awaits a more explicit vision of the most productive foci on which to concentrate the powers of the limited personnel. This is required soon because the available personnel and financial resources are growing fast. The foci of the research of PhD students have thus far been relatively widespread (although this may partly be due to the fact that many PhD students started their project in the programme leaders' previous research programme at Maastricht University). This makes it difficult to deduce a clear vision or set of strategies with regard to where the unit is going. While growing quickly and successfully, the programme is still relatively small, and in order to create 'ground-breaking' results of international dimension and to have important societal value, it will have to focus its efforts and take more time to fulfill the optimistic promises.

Appendices

Appendix A: Explanation of the SEP criteria and scores

The four main criteria for assessment are: Quality, Productivity, Relevance, and Vitality & Feasibility. The assessment at the institute level primarily focuses on strategy and organisation, whereas the assessment at the level of the research group or programme primarily focuses on the performance and activities of researchers and the results of their work (output and outcome).

Quality	The level or degree of excellence of the research, compared to accepted (international) standards in that field. The scope of the term ‘research’ is not limited to the research results. Research management, research policy, research facilities, PhD training and the societal relevance of research are considered integral parts of the quality of work in an institute and its programmes.
Productivity	The relationship between input and output, judged in relation to the mission and resources of the institute.
Relevance	Social, economic and cultural relevance. Aspects to be considered are: <ul style="list-style-type: none"> • <i>Social quality</i>: efforts of the institute or group to interact in a productive way with stakeholders in society • <i>Social impact</i>: how research affects specific stakeholders or procedures in society • <i>Valorisation</i>: activities aimed at making research results available and suitable for application in product, processes and services. Committee members can also remark on the relevance for the academic community, but the assessment should be on societal relevance.
Vitality & Feasibility	The ability to react adequately to important changes in the environment. Also vision for the future.

The scores on a five-point scale are:

5 Excellent	Research is world leading. Researchers are working at the forefront of their field internationally and their research has an important and substantial impact in the field.
4 Very Good	Research is considered nationally leading. Research is internationally competitive and makes a significant contribution to the field.
3 Good	Research is considered internationally visible. Work is competitive at the national level and makes a valuable contribution in the international field.
2 Satisfactory	Research is nationally visible. Work adds to our understanding and is solid, but not exciting.
1 Unsatisfactory	Work is neither solid nor exciting, flawed in the scientific and/or technical approach, repetitions of other work, etc.

Appendix B: Programme for the site visit

Schedule research review Pedagogics, Utrecht, 23-28 March 2013

Saturday	March	23		Hotel Mary K, 'werfkelder' (meeting room)	
	15.00	16.00		introduction: context and procedures of research reviews in the Netherlands	Committee & Secretary Prof. Paul Vedder (LEI)
	16.00	18.00		preparatory committee meeting	Committee & Secretary
Sunday	March	24		Hotel Mary K, 'werfkelder' (meeting room)	
	9.30	10.00		committee meeting	Committee & Secretary
	10.00	11.00	VU	meeting with institute management	<ul style="list-style-type: none"> • Prof. Jan Passchier (dean) • Prof. Carlo Schuengel (associate dean research)
	11.15	12.00		Challenges to childrearing relationships	<ul style="list-style-type: none"> • Prof. Carlo Schuengel • Prof. Catrin Finkenauer
	12.15	13.00		Education for the good life in a diverse society	<ul style="list-style-type: none"> • Prof. Doret de Ruyter • Prof. Bert van Oers
	13.15	14.15		Lunch break	Committee & Secretary
	14.15	15.00		Educational neuroscience	<ul style="list-style-type: none"> • Prof. Lydia Krabbendam • Prof. Jelle Jolles
	15.15	16.00		meeting with PhD students	<ul style="list-style-type: none"> • Marije Verhage • Asuman Buyukcan Tetik • Chiel van der Veen • Jacomijn van der Kooij • Sanne Dekker • Iro Xenidou-Dervou
	16.15	17.15		committee meeting	Committee & Secretary

Monday	March	25		Academiegebouw, Kanunnikenzaal/Sterrecamer	
	8.30	9.30	UvA	meeting with institute management	<ul style="list-style-type: none"> • Prof. E.H.F. de Haan (dean) • Prof. F.J. Oort (director research institute)
	9.30	10.15		Child development	<ul style="list-style-type: none"> • Prof. G.J.J.M. Stams • Prof. S.M. Bögels • Prof. J.M.A. Hermanns
	10.30	11.15		Education	<ul style="list-style-type: none"> • Prof. M.L.L. Volman • Prof. P.F. de Jong
	11.30	12.15		meeting with PhD students	<ul style="list-style-type: none"> • Drs. K.O.W. Helmerhorst • E.L. Möller MSc • A. van Maanen MSc • M. van den Boer MSc • T.E. Hornstra, MSc • B.B.F. Eidhof MSc
	12.30	13.30		lunch break	Committee & Secretary
	13.30	14.30	RUG	meeting with management	<ul style="list-style-type: none"> • Prof. Henk Kiers (dean) • Prof. Roel Bosker (director research)
	14.45	15.30		Education in Culture	<ul style="list-style-type: none"> • Prof. Jeroen Dekker • Prof. Greetje Timmerman
	15.45	16.30		Research and Evaluation of Educational Effectiveness	<ul style="list-style-type: none"> • Prof. Greetje van der Werf • Prof. Roel Bosker
	16.45	17.30		Developmental and Behavioral	<ul style="list-style-type: none"> • Prof. Hans Grietens • Prof. Alexander Minnaert

	17.30	18.15		meeting with PhD students	<ul style="list-style-type: none"> • Drs. Marieke Dekker (EiC) • Linda Greveling MSc (EiC) • Mayra Mascareño MSc (RaEoEE) • Mechteld van Kuijk MSc (RaEoEE) • Marlous Tiekstra MSc (DaB) • Els Evenboer MSc (DaB)
	18.15	19.00		committee meeting	Committee & Secretary
Tuesday	March	26		Academiegebouw, Kanunnikenzaal/Sterrecamer	
	8:30	9:30		committee meeting	Committee & Secretary
	9.30	10.30	LEI	meeting with institute management	<ul style="list-style-type: none"> • Prof. Hanna Swaab (dean) • Prof. Judi Mesman (academic director Pedagogics)
	10.30	11.15		Child and family	<ul style="list-style-type: none"> • Prof. Marian Bakermans-Kranenburg • Prof. Femmie Juffer • Prof. Marinus van IJzendoorn
	11.30	12.15		Child and educational settings	<ul style="list-style-type: none"> • Prof. Adriana Bus • Prof. Paul van den Broek
	12.30	13.30		lunch break	Committee & Secretary
	13.30	14.15		Child welfare services	<ul style="list-style-type: none"> • Prof. Hanna Swaab • Prof. Paul Vedder

	14.30	15.15		meeting with PhD students	<ul style="list-style-type: none"> • Anne Helder (Educational Settings) • Anouk Goemans (Welfare Services) • Thijs Nielen (Educational Settings) • Mariëlle Prevoo (Family) • Madelon Riem (Family) • Jill Suurland (Welfare Services)
	15.30	16.30		committee meeting	Committee & Secretary
Wednesday	March	27		Academiegebouw, Kanunnikenzaal en Sterrecamer	
	8.30	9.30	UU	meeting with institute management	<ul style="list-style-type: none"> • Prof. Werner Raub (dean) • Prof. Anneloes van Baar (coordinator research assessment)
	9.30	10.15		Education and Learning	<ul style="list-style-type: none"> • Prof. Marian Jongmans • Prof. Mieke Brekelmans • Dr. Sanne Akkerman • Prof. Paul Leseman
	10.30	11.15		Child and adolescent Studies	<ul style="list-style-type: none"> • Prof. Wilma Vollebergh • Prof. Maja Deković • Prof. Susan Branje • Prof. Anneloes van Baar

	11.30	12.15		meeting with PhD students	<ul style="list-style-type: none"> • Larike Bronkhorst (EL) • Ora Oudgenoeg-Paz (EL) • Emilie Prast (EL) • Margot Peeters (CAS) • Stephanie Nelemans (CAS) • Daphne van de Bongardt (CAS)
	12.30	13.30		lunch break	Committee & Secretary
	13.30	14.30	UT	meeting with management	<ul style="list-style-type: none"> • Prof. Cees Aarts (academic director) • Prof. Karen van Oudenhoven-van der Zee (dean)
	14.30	15.30		Educational design and evaluation	<ul style="list-style-type: none"> • Prof. Peter Slegers • Prof. Ton de Jong • Prof. Cees Glas
	15.30	16.15		meeting with PhD students	<ul style="list-style-type: none"> • Marjolein van Klink Msc • Frank Leenaars Msc • Tjark Huizinga Msc • Arnoud Oude Groote Beverborg Msc • Dr. Hanneke Geerlings
	16.30	17.30		committee meeting	Committee & Secretary
Thursday	March	28		Hotel Mary K, 'werfkelder' (meeting room)	
	8.30	12.00		committee meeting, writing	Committee & Secretary

Appendix C: Short Curriculum Vitae of the committee members

Lieven Verschaffel obtained in 1984 the degree of Doctor in Educational Sciences at the University of Leuven, Belgium. From 1979 until 2000 he fulfilled several research positions at the Fund for Scientific Research Flanders. Since 2000 he is a full professor in educational sciences of that same university, and since 2003 he is the head of the Center for Instructional Psychology and Technology. His major research interests are educational psychology and (psychology of) mathematics education. Lieven Verschaffel is a member of the editorial board of numerous international journals, including *Educational Studies in Mathematics*, *Mathematical Thinking and Learning*, *Educational Research Review*, *Learning and Instruction*, *Human Development*, and *Cambridge Journal of Education*. He is Series Editor (together with Wolff-Michael Roth) of the book series *New Directions in Mathematics and Science Education* published by Sense Publishers. He has published extensively about his research and he has been asked to give plenary talks at several major international conferences in his area of research. For his contribution to (mathematics) education, he has been honored several times. In 2009, he was elected in 2009 as Member of the Royal Academia of Belgium for Sciences and Arts, and, in 2010 as a Member of the Academia Europae.

Mary Dozier Mary Dozier is Amy E. du Pont Chair of Child Development and Professor of Psychology at the University of Delaware. She studies the development of children who have experienced early adversity, and the effectiveness of parenting programmes for such children. She is currently conducting randomised clinical trials with high-risk birth children, foster children, and internationally adopted children, work that has been supported by the National Institute of Mental Health continuously since 1989. She is on the Institute of Medicine's Committee on Child Maltreatment, is a previous associate editor of *Child Development*, and serves on a number of advisory and editorial boards.

Mark Greenberg holds The Bennett Endowed Chair in Prevention Research in Penn State's College of Health and Human Development where he has taught since 1997. He is the Founding Director of the Prevention Research Center for the Promotion of Human Development. He is the author of more than 300 journal articles and book chapters on developmental psychopathology, well-being, and the effects of prevention efforts on children and families. He is a senior investigator on a series of large-scale multi-site studies that focus on both the prevention of substance abuse and serious conduct disorder, the promotion of social and emotional well-being and the the implementation and sustainability of community-based prevention, and the promotion of mindfulness. One of his current interests is how to help nurture awareness and compassion in our society.

Eckhard Klieme is Professor of Educational Science at the Johann Wolfgang Goethe University in Frankfurt am Main. He studied mathematics, psychology, communication studies and education science at the University of Bonn. There he received his doctorate in psychology in 1988 and taught from 1986 to 1989 as a lecturer at the Department of Psychology. From 1983 to 1997 he was a research associate at the Institut für Test- und Begabungsforschung of the German National Academic Foundation in Bonn. From 1998 to 2001 he was a research associate at the Max Planck Institute for Educational Research, with the research domains of educational science and education systems. Since 2001, Klieme is Professor of Educational Science at the Goethe University. He also heads the Working Group "Education Quality and Evaluation" at the German Institute for International Educational Research (DIPF) and was its director from 2004 to 2008.

Heikki Lyytinen, Professor of developmental Neuropsychology, led EU-COST A8 “Learning Disorders as a Barrier to Human Development” action from 1994-1998, co-lead with Lea Pulkkinen the Centers of Excellence “Human Development and its risk factors” 1997-2006 and with Jari-Erik Nurmi “Learning and Motivation” 2006-2011, both funded by the Academy of Finland. He has also directed the Jyväskylä Longitudinal study of Dyslexia (JLD) from 1993. He was Vice President of the UJ 1997-2000. Today he is the Chair of the Boards of the Agora Human Technology Centre of UJ and that of Niilo Mäki Foundation maintaining the Niilo Mäki Institute (www.nmi.fi/front-page). He is a member of the Academy of Sciences and Letters (of Finland, 2003-). He has published extensively in international journals in his main domains of interest, viz. psychophysiology/ cognitive neuroscience, dyslexia and reading acquisition (for publications see heikki.lyytinen.info) and has most recently focused his research on preventive training for children at risk of reading difficulties or dyslexia (see grapholearn.info).

Neil Mercer is Professor of Education (2000) at the University of Cambridge, where he is also Chair of the Psychology and Education Group and Vice-President of the college Hughes Hall. Previously, he was Professor of Language and Communications at the Open University. He is a psychologist with particular interests in the relationship between language and thinking, the development of children’s language and reasoning, the role of the teacher in guiding learning and the educational functions of classroom talk. His research with colleagues generated the *Thinking Together* practical approach to classroom pedagogy, and he has worked extensively with teachers, researchers and educational policy makers on improving talk for learning in schools. Formerly editor of the journals *Learning and Instruction* and *International Journal of Educational Research*, he is now an editor of *Learning, Culture and Social Interaction*. His most recent books are *Exploring Talk in School* (with Steve Hodgkinson), *Dialogue and the Development of Children’s Thinking* and *Intertinking: putting talk to work* (both with Karen Littleton).

Karine Verschueren obtained her PhD in 1996 from the University of Leuven, Belgium in the domain of developmental psychology, on the topic of attachment, self-concept, and socio-emotional functioning in early childhood. In 2000 she was appointed as professor in the domain of school psychology at the University of Leuven. At this moment, she is full professor at the research unit School Psychology and Child and Adolescent Development, where she studies the development of children and adolescents in the school context and the risk and protective factors at the level of child, family, and school. Contextual factors considered primarily refer to proximal influences, including interpersonal relationships with parents, teachers, and peers. Aspects of child adjustment most frequently considered include externalizing behavior, self-concept, academic engagement, and achievement. In addition she studies students’ school trajectories and their adjustment to school transitions, using longitudinal research designs. She is member of the editorial board of *Attachment and Human Development* and *Journal of School Psychology*.

Appendix D: Citation-analysis by Ad Prins

Introductory briefing

This bibliometric analysis intends to provide an international comparison for research programs in Education and Pedagogical Sciences answering the question “How does the impact of the Dutch educational and pedagogical scientists compare to the impact of their international colleagues?”. The analysis is based on a set of the output of programs in the fields of Education and Pedagogical Sciences in the Netherlands. Programs were asked to send in 10 publications for each year to be evaluated between 2006 and 2011, resulting in 60 publications per program, in one case 30.

The design of the analysis addresses the question of international comparison, accounting for the academic diversity among programs in the field. Disciplinary differences among programs are manifest in publication and citation culture. The output of programs results in a wide range of journals and publication formats (including also scientific reports, books and chapters in volumes). Because of this, Google Scholar has been chosen as a more suitable data source over others, such as Web of Science or Scopus. The results of the analysis are presented per program. Results from the analysis are not suited for mutual comparison of programs.

To achieve comparison while addressing the issue of disciplinary diversity, three comparisons have been chosen. The comparisons consist of various ways to define disciplinary domains of the fields. For each domain, quotients are computed that represent the (international) average *impact of the domain*. The quotients are based on totals of citations since publication divided by the number of documents published. Data for impact quotients and domains are derived from Scopus.¹ These data are compared with the *impact of programs*, which is computed as the total of citations of the publications for each program divided by the number of publications involved.

The three comparisons are:

A. *Group Specific Domains*: The basis of comparison in this approach is formed by the six journals in which each program publishes most often. Averages of the citations of articles of programs published in the *six journals of each program* are compared with the average impacts of these journals.

B. *Discipline Specific Domains*: The average impact of each program is compared with the average impact of journals in two domains that relate to the disciplinary differences in the field; *Education* and *Pedagogical Sciences*. The journals are part of a long list of 457 journals taken from of several Scopus science categories relevant to these domains. Each journal in the list has been categorized by researchers in the field to represent either the domain of Education (322 journals) or Pedagogical Sciences (118), or both (17). As an indication for the relevance of the two domains for the program, a percentage is given of how many of its submitted publications are in journals of either domain.

C. *General Scopus Domains*: The impact of each program is compared with the impact of journals in three general Scopus domains. The relevant Scopus domains are: *Education*, *Psychology Miscellaneous*, and *Psychiatry & Mental Health*. As an indication for the relevance of the

¹ For an extensive description of definitions, methods and quality of data, See Addendum 1.

Scopus science category for the program, a percentage is given of how many of its submitted publications are in journals of this category.

Programs differ considerably in size, which could not be accounted for. Whereas all groups added the same number of publications to the total pool of publications, the selection base within programs depends on the size of the program. The number of highly visible publications to choose from might have been limited in the smaller programs, resulting also in lower impact averages. The results of the analyses are therefore not suited for comparison of programs across the board, but rather allow a comparison of impact of selected publications of the programs with international standards.

Legend

Impact of programs

- The impact of programs is computed as (1) the average of citations of the publications published in year X (2006 or later), and (2, in italics) in case of Overall Averages or Average of Averages, the sum of all citations since publication date divided by the total number of these publications.
- Comparison A is presented in a table separate from comparisons B and C because the impact of programs in comparison A is based on a selection of the publications (articles in six journals in which authors of the program publish most often).

Impact of journals and domains

- The impact of a domain (according to comparison A, B or C) is (1) the average of impacts of all journals in the domain for a particular year, and (2, in italics) the average of all averaged journal impacts in that domain. The impact of a journal in a particular year X is the average of all citations of publications in the journal over the period of three years (year X and two previous years X-1 and X-2) divided by the number of documents published in that period. Scopus data are used to compute journal impacts for each domain. The comparison with Scholar data is enabled by transforming these Scopus data by a formula derived from Non-linear multilevel regression analysis. See addendum 1.

Graphs: Boxplots

As averages in the tables for comparisons A, B and C may suggest a level of exactness that defies the variations in the underlying data (both for programs and for the journal data), graphs represent the distribution of these data with boxplots, representing the minimum, first quartile, *median* (thick horizontal line somewhere in the middle of the box), third quartile, and maximum of the data. Also, extreme values and outliers are represented with 0 and * respectively. The numbers of extreme and outlier values refer to IDs of individual cases of journals.² The Y-axes of the graphs (Citationvalue) represent numbers of citations, on a scale (normal or log scale) that fit the results of programs.

Comparison A. Group Specific Domains

- *Articles Involved*: number of articles in the journals most often used by the program, which are the basis for calculating the impact of the program. Is the number of articles relatively

² ID numbers of journals may vary per graph, due to deletion of missing values.

low in comparison with the total number of publications (see below), the domain impact might be regarded to be less relevant as a norm for the program.

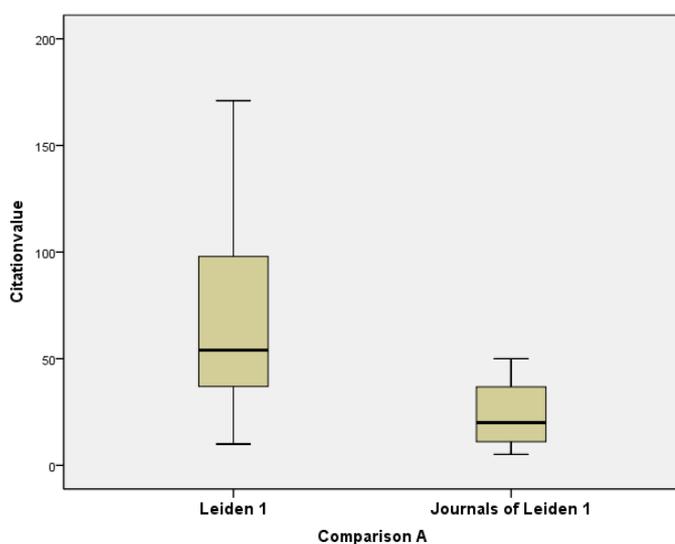
Comparisons B (Discipline Specific Domains) and C (General Scopus Domains).

- *Coverage:* As an indication for the relevance of the domain for the program a percentage is given of how many of its submitted publications are in journals of the specified domain. Is the percentage low, the domain impact is to be regarded to be less relevant as a norm for the program.
- *Number of publications:* number of publications (articles, books or chapters in volumes) included in the analysis.

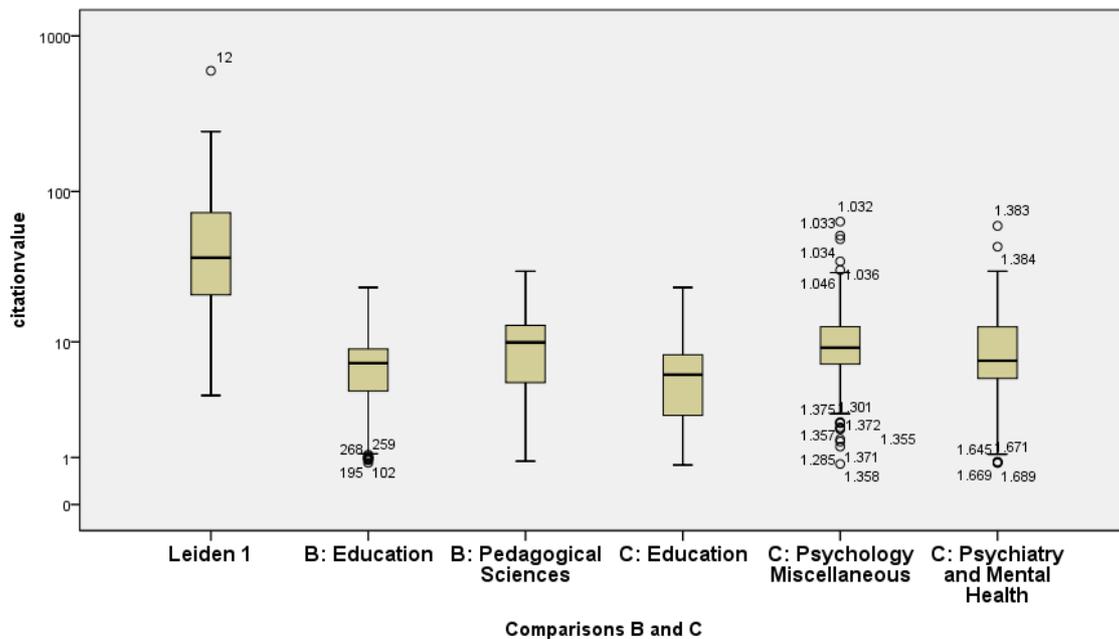
Results

Bibliometric results for Leiden 1

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
Leiden 1	99	72,5	57,75	48,33	40,25	68	68,68	22
Journals of Leiden 1	38,72	40,01	20,81	11,04	11,23	5,15	23,76	

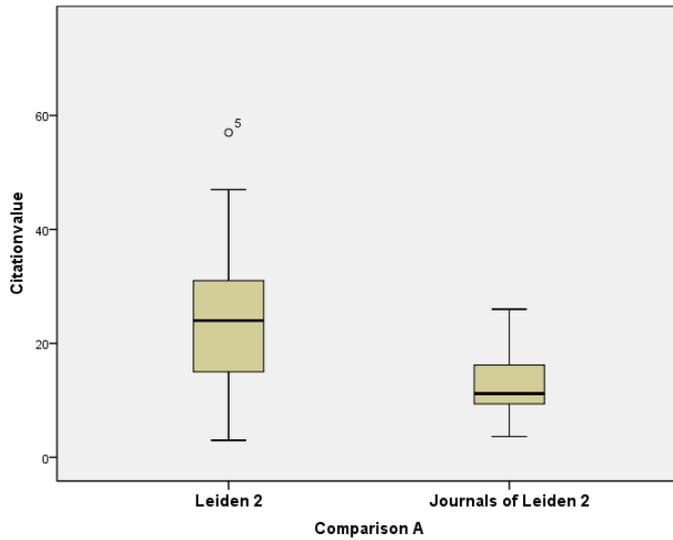


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av of Averages	Coverage
Leiden 1	93,4	121,3	65,89	31,4	29,6	23,8	60,81	59 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	0%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	46%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	9%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	32%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	21

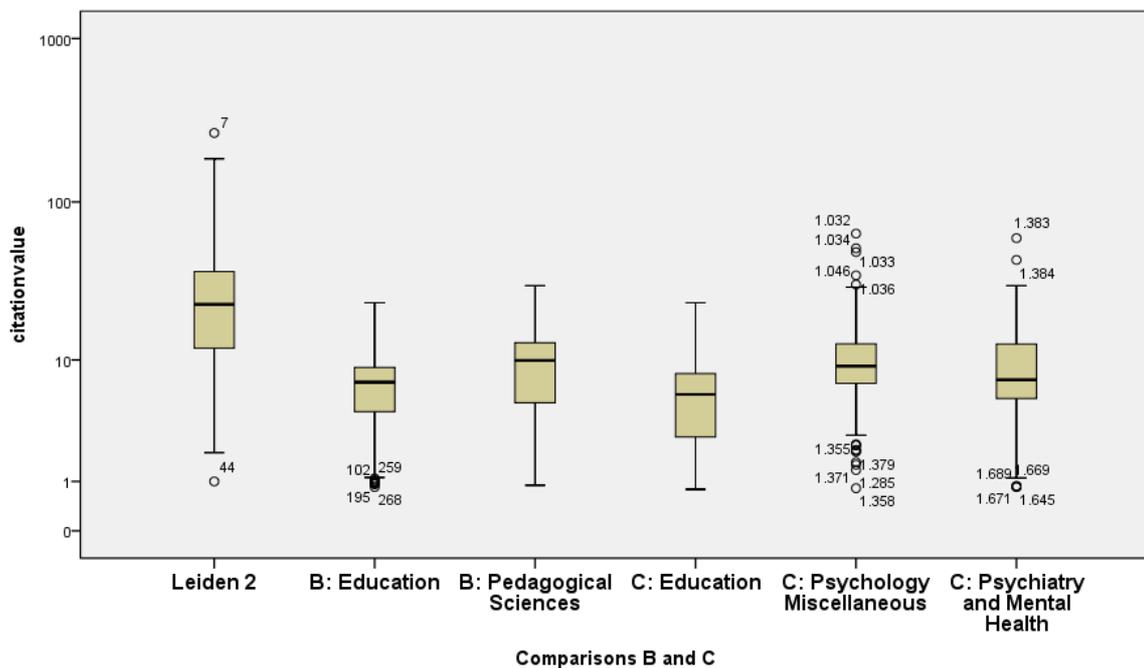


Bibliometric Results for Leiden 2

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
Leiden 2	265	38,5	23	30,75	16,14	3	36,24	21
Journals of Leiden 2	26,01	18,69	14,63	11,99	9,33	3,64	12,46	

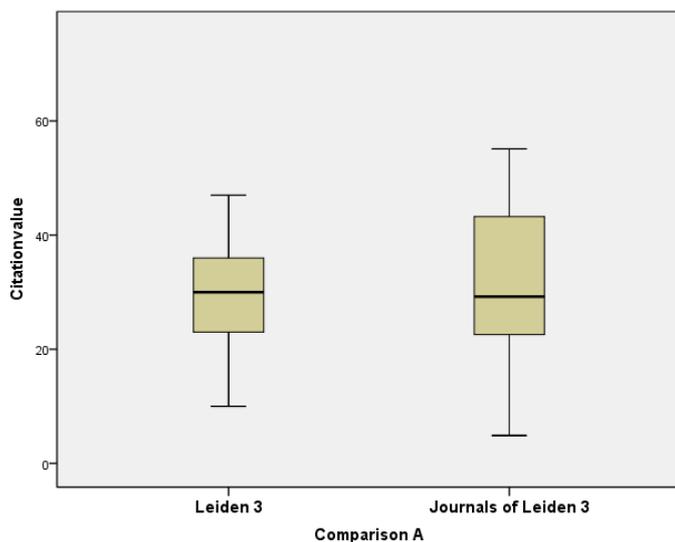


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
Leiden 2	95,89	41,7	27,3	25,33	14,7	6,89	34,91	59 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	75%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	0%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	53%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	22%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	0%

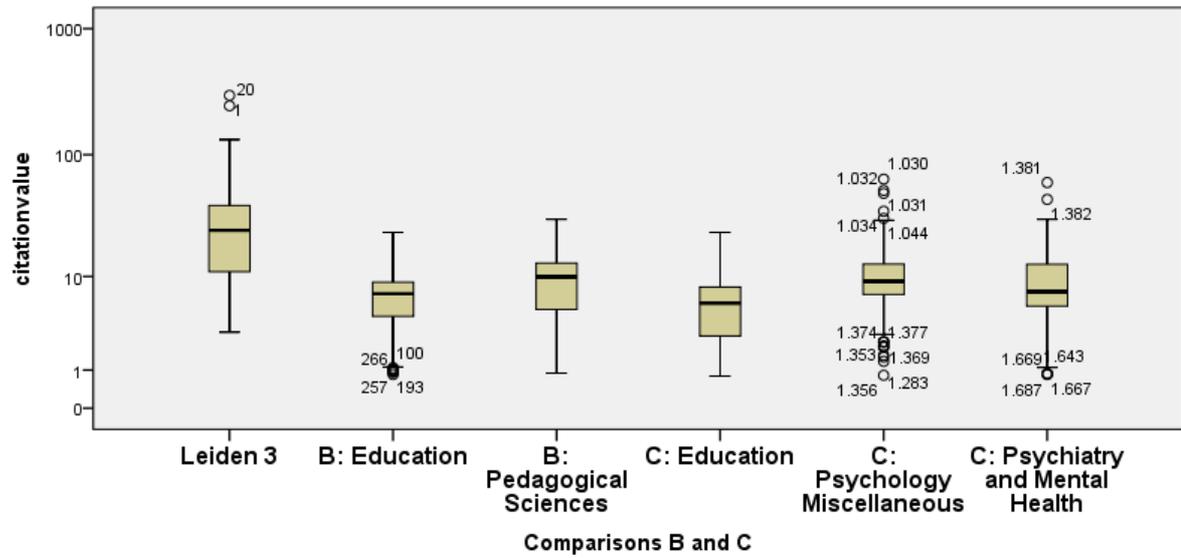


Bibliometric Results for Leiden 3

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
Leiden 3	89,5	29,5	32,8	27	15,5	10	<i>34,17</i>	18
Journals of Leiden 3	51,27	40,43	28,4	25,69	14,12	4,87	<i>31,76</i>	

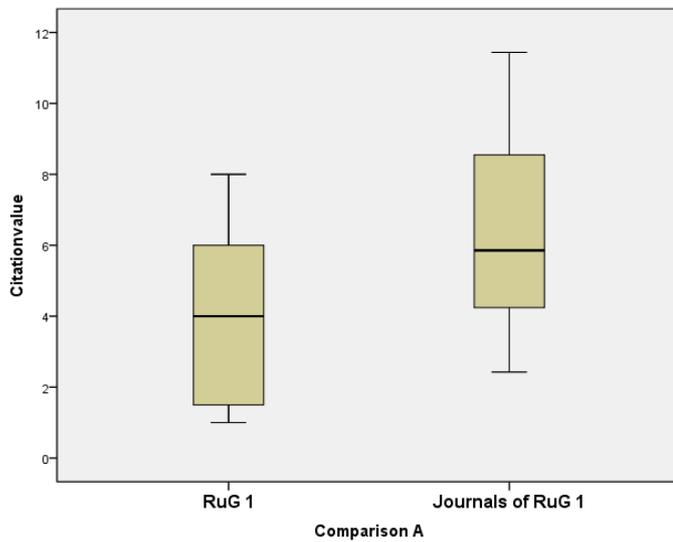


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
Leiden 3	79,2	35,11	52,4	31,78	12,8	6,2	<i>36,34</i>	58 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	<i>8,42</i>	0%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	<i>11,5</i>	22%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	<i>7,83</i>	0%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	<i>11,49</i>	18%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	<i>11,69</i>	29%

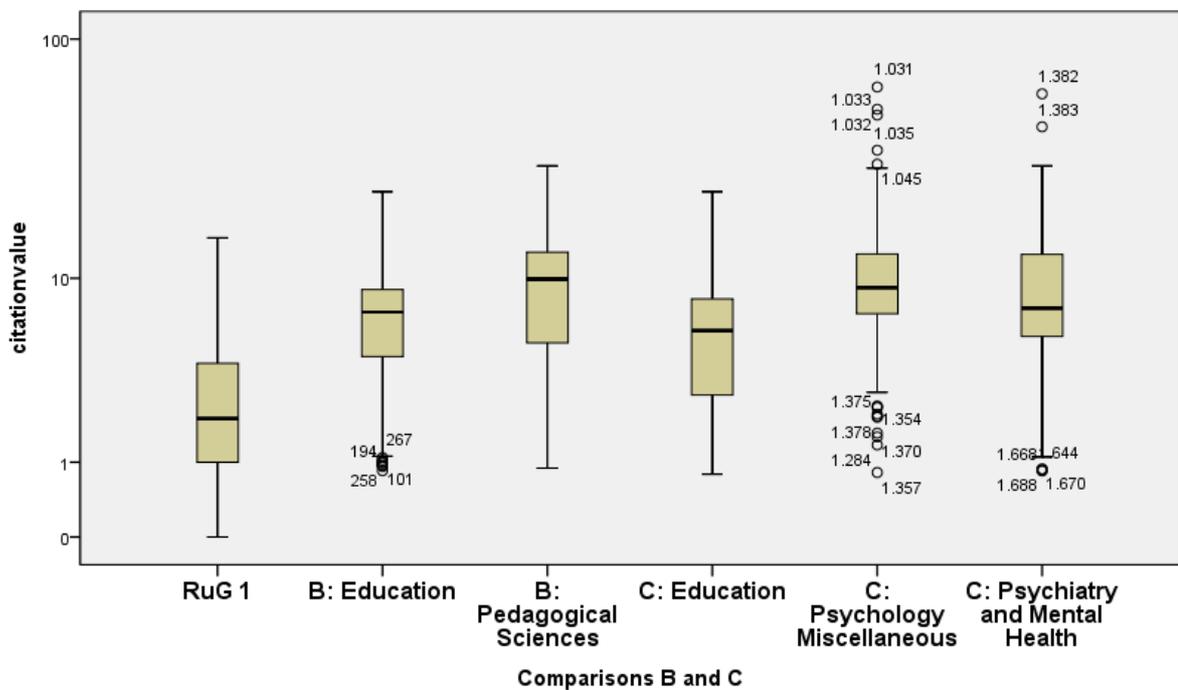


Bibliometric Results for RUG 1

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
RuG 1	n.d.	4	2,5	6,67	1,5	n.d.	3,5	15
Journals of RuG 1	n.d.	9,55	7,44	5,41	2,54	n.d.	5,57	

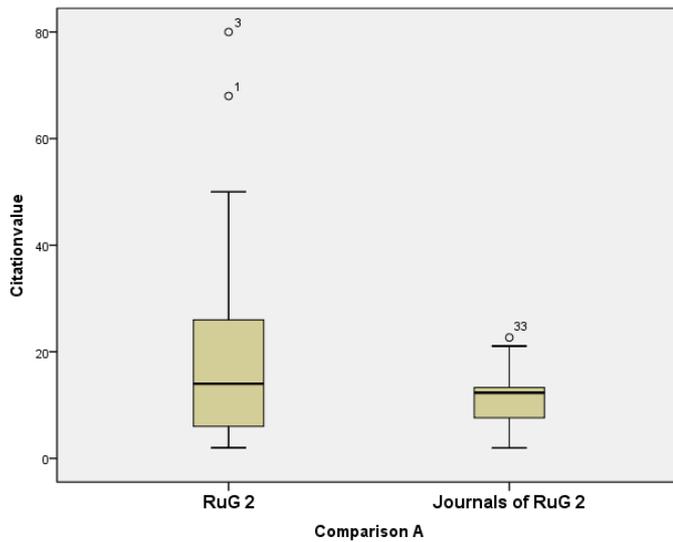


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
RuG 1	5,13	3,33	2,3	4,29	2,5	1,67	3,31	59 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	24%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	30%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	54%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	5%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	0%

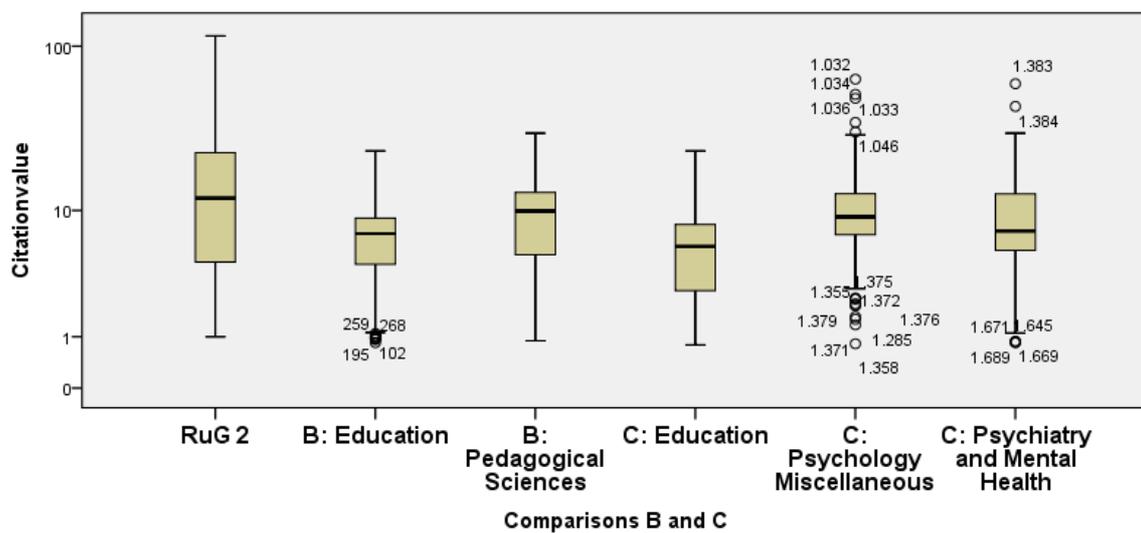


Bibliometric Results for RUG 2

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
RuG 2	46,4	20,83	17,5	12,5	14,33	5	21	25
Journals of RuG 2	16,61	15,38	12,48	9,45	7,4	2,24	11,44	

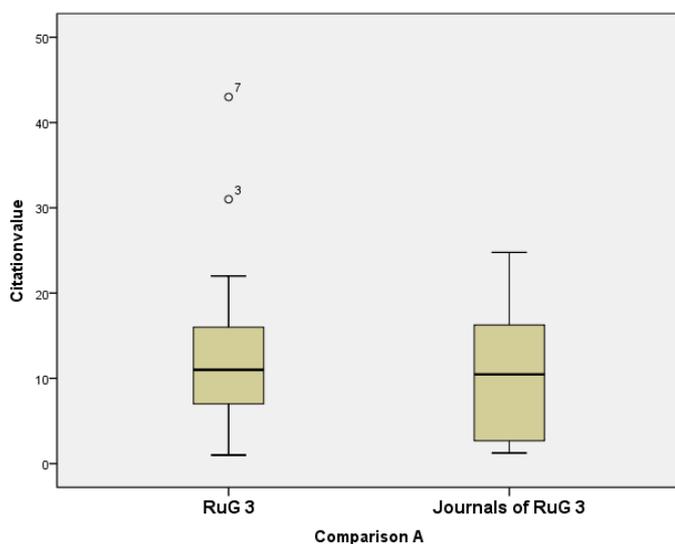


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
RuG 2	34,4	23,1	27	10,9	7,9	3,56	18,05	59 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	66%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	5%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	60%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	16%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	0%

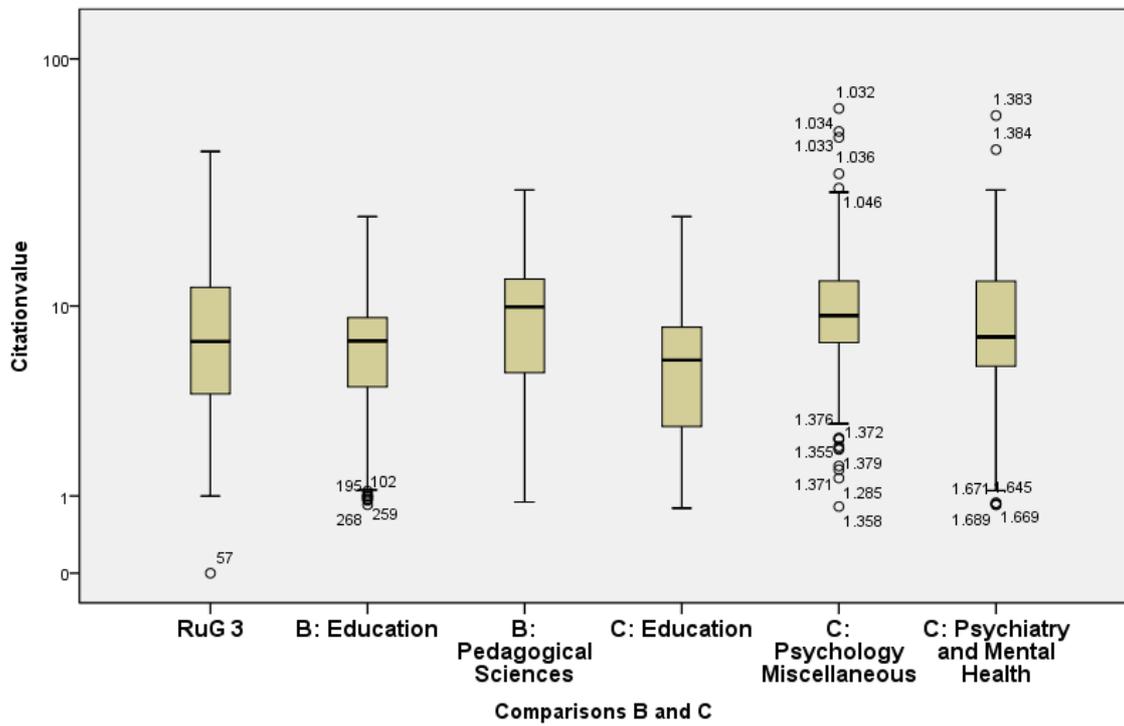


Bibliometric Results for RUG 3

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
RuG 3	11	12,5	28	12	10	5	<i>12,67</i>	19
Journals of RuG 3	13,55	18,02	14,61	13,7	6,45	2,13	<i>8,27</i>	

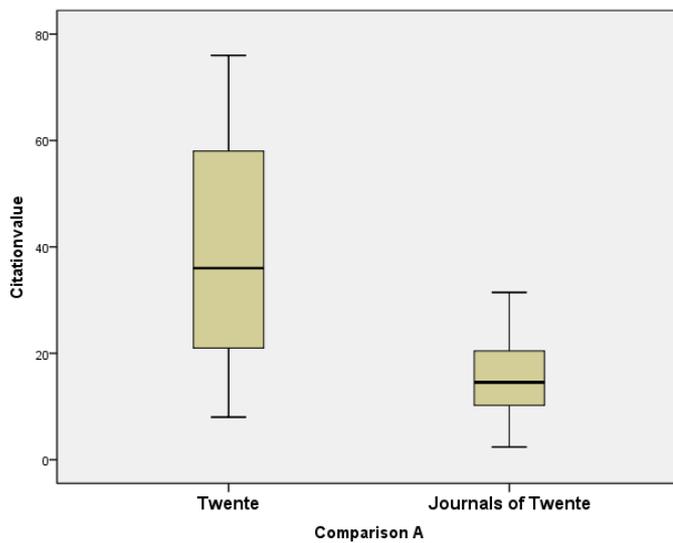


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
RuG 3	10,29	12,2	13,1	8,2	8,5	4,44	<i>9,5</i>	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	<i>8,42</i>	13%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	<i>11,5</i>	13%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	<i>7,83</i>	17%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	<i>11,49</i>	15%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	<i>11,69</i>	17%

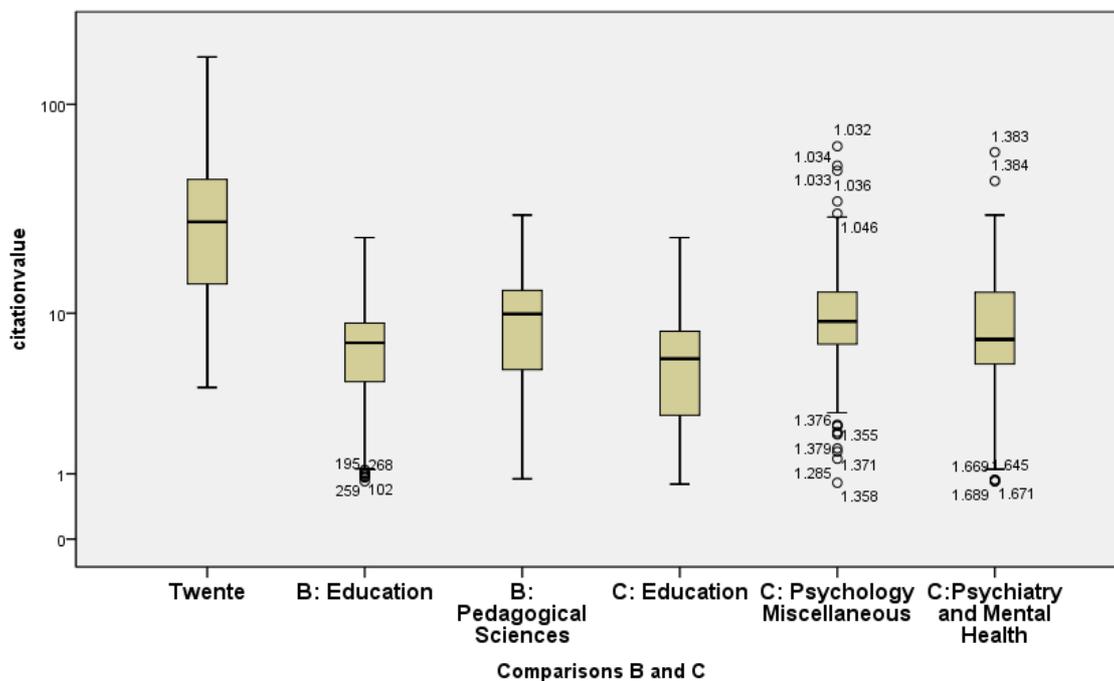


Bibliometric Results for Twente

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
Twente	68,67	49,33	37	28,5	15,5	9	37,94	17
Journals of Twente	24,83	19,46	15,66	9,57	8,3	3,22	14,9	

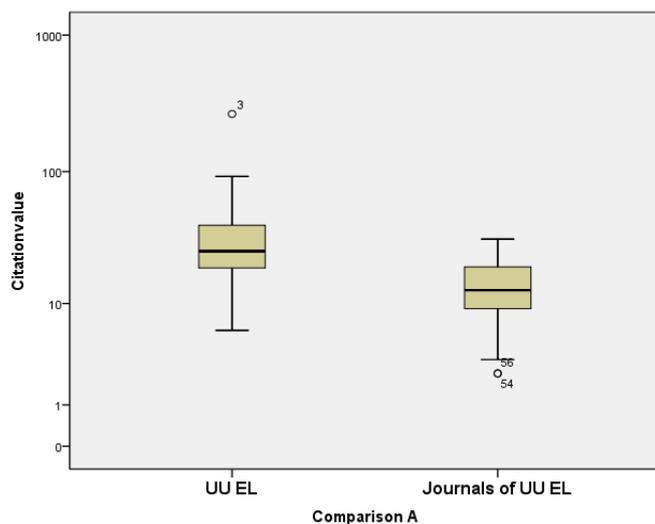


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
Twente	68,7	44,5	38,6	20,1	17,6	7,9	32,9	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	56%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	0%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	44%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	14%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	2%

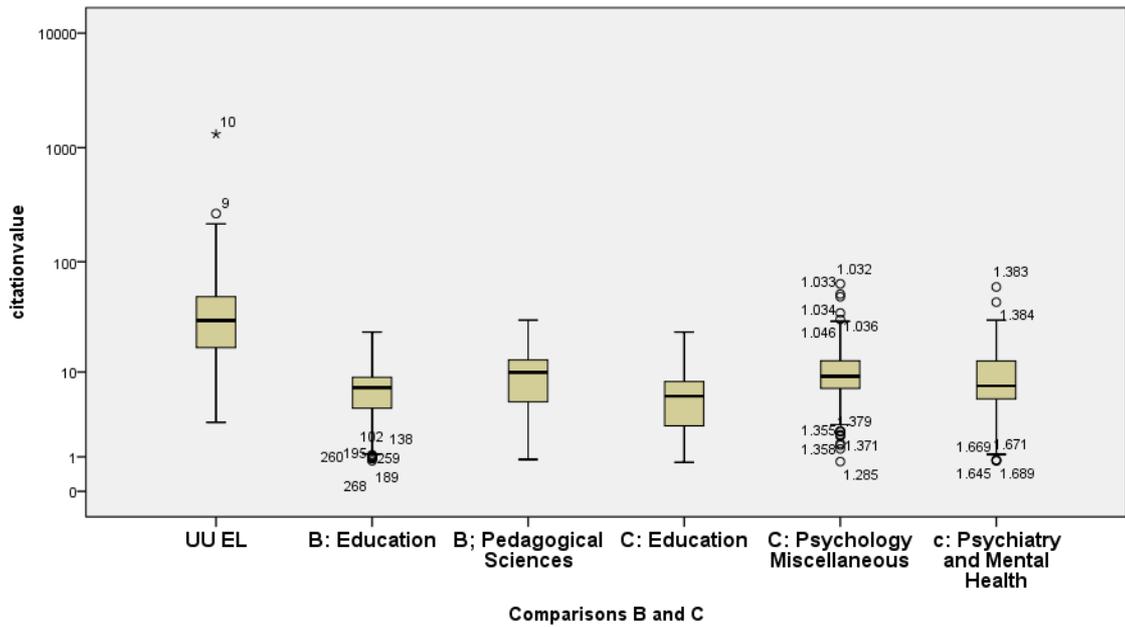


Bibliometric Results for UU EL

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
UU EL	116,67	57,2	25,33	22,5	21,4	11	37,96	28
Journals of UU EL	25,66	22,12	14,66	10,07	8,75	2,69	13,85	

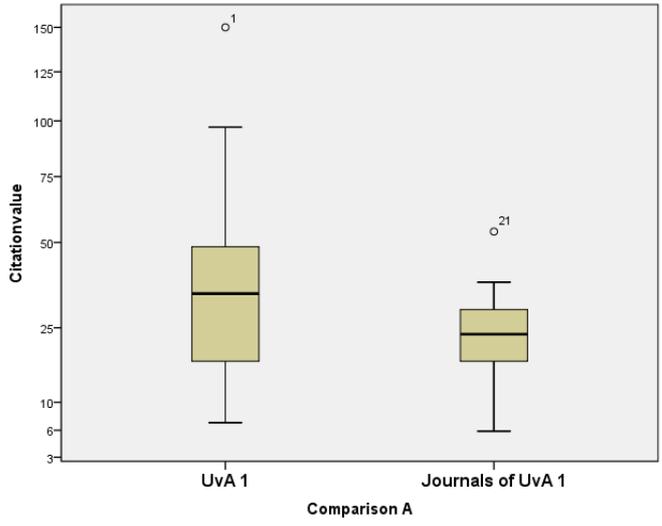


Comparison B and C	2006	2007	2008	2009	2010	2011	Average of Averages	Coverage
UU EL	215	74,7	31,2	29	20,5	10,4	63,47	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	71%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	7%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	50%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	12%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	2%

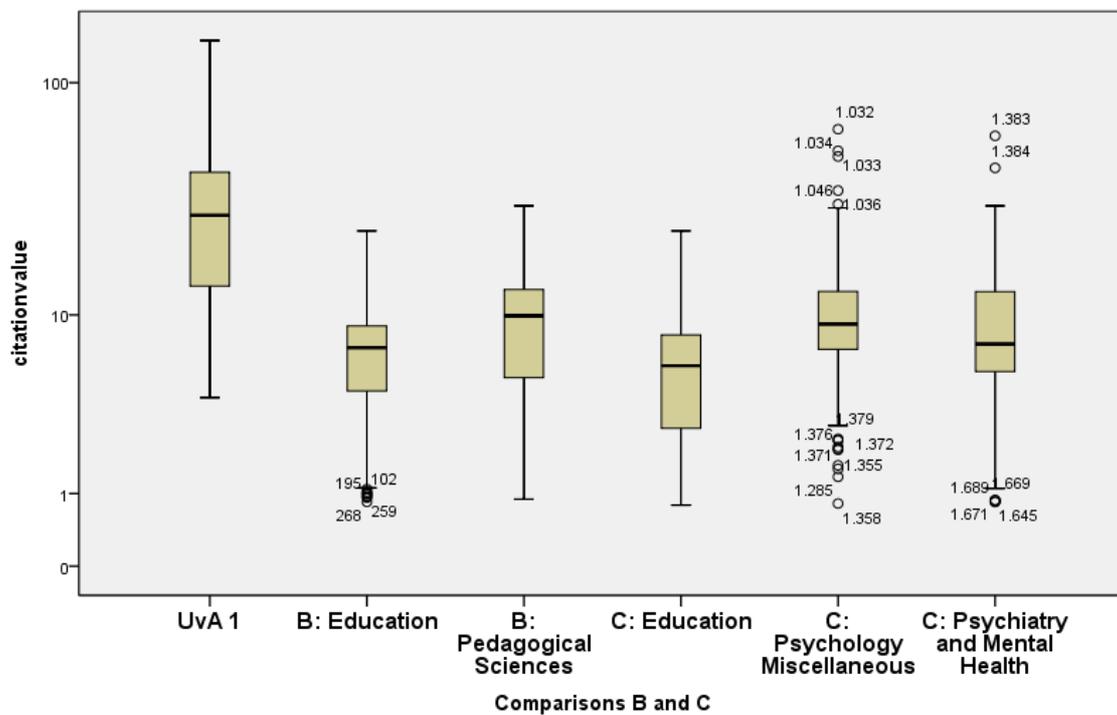


Bibliometric Results for UvA 1

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
UvA 1	70,4	40,5	53,5	26,4	15,67	7	41,65	20
Journals of UvA 1	33,33	28,82	28,96	20,49	11,19	6,56	24,14	

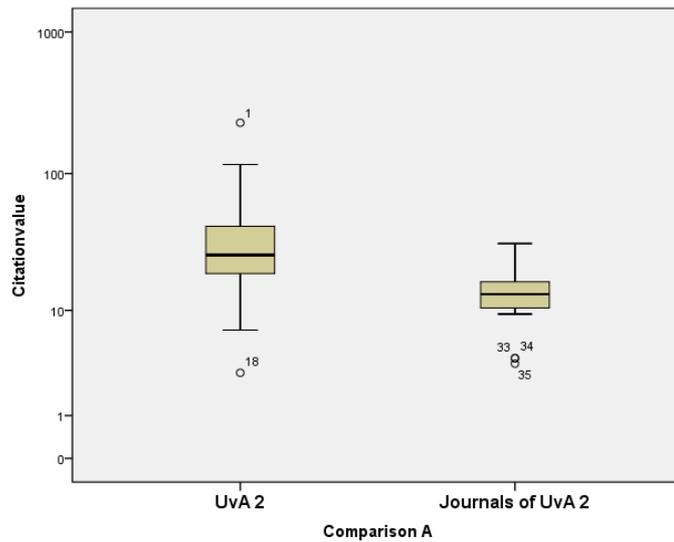


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
UvA 1	51,5	34,1	45,2	21	24,8	7,5	30,68	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	0%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	29%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	2%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	31%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	29%

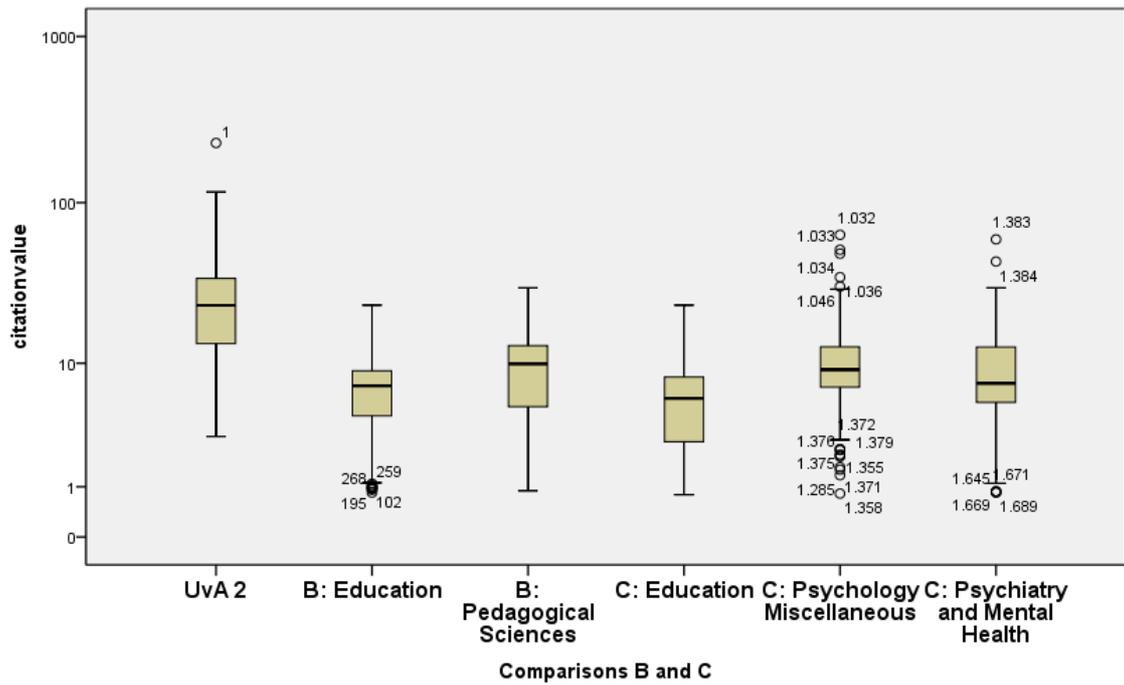


Bibliometric Results for UvA 2

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
UvA 2	70,67	38,5	30	22,5	20	6	42,83	18
Journals of UvA 2	20,77	19,68	14,92	13,29	10,33	3,92	13,07	

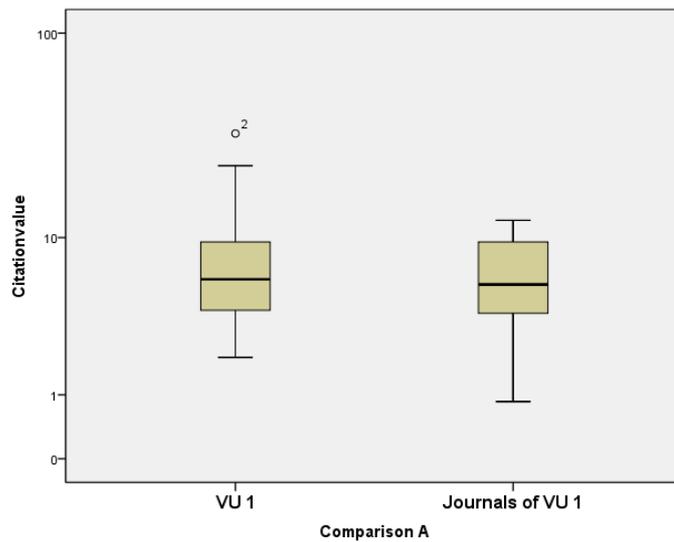


Comparison B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
UvA 2	67,9	41,7	27,4	22	15,1	8,1	30,37	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	68%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	2%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	63%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	16%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	0%

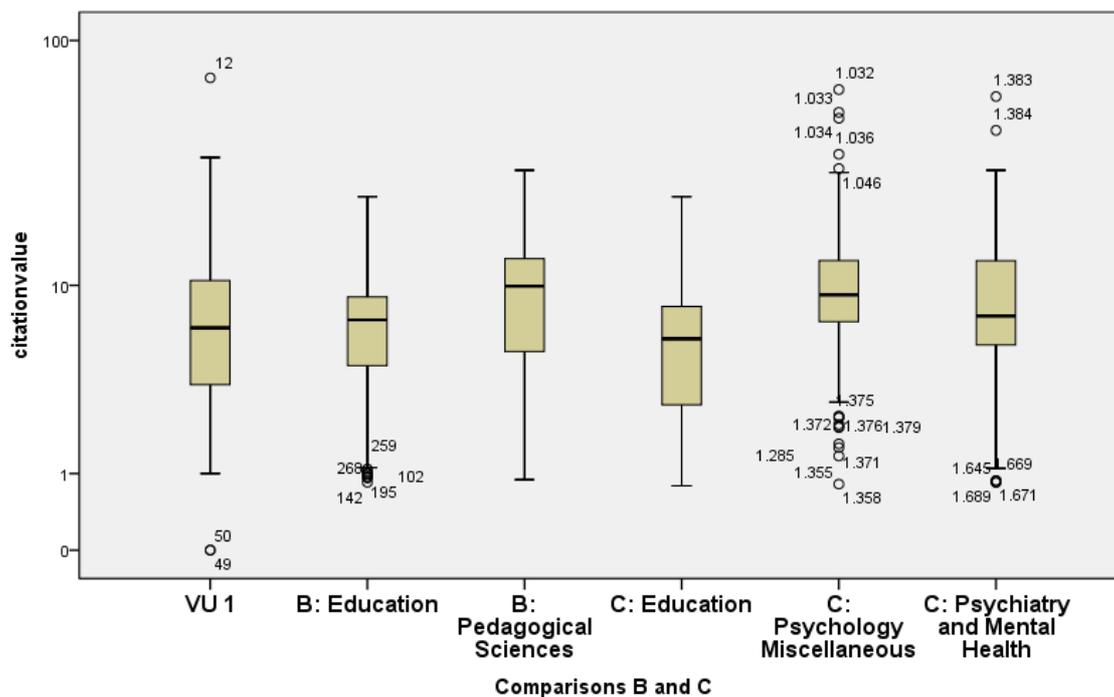


Bibliometric Results VU 1

Comparison A	2006	2007	2008	2009	2010	2011	Avarage of Averages	Articles involved
VU 1	16	11	12,33	6	7	2	8,93	15
Journals of VU 1	11,54	11,47	7,27	4,81	4,59	1,49	6,4	

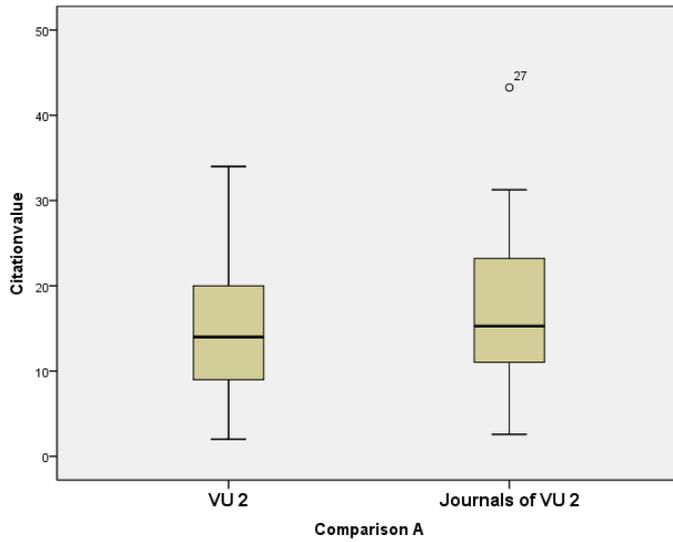


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
VU 1	9,2	21,2	13,3	7,2	5	1,9	9,63	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	65%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	2%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	70%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	0%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	0%

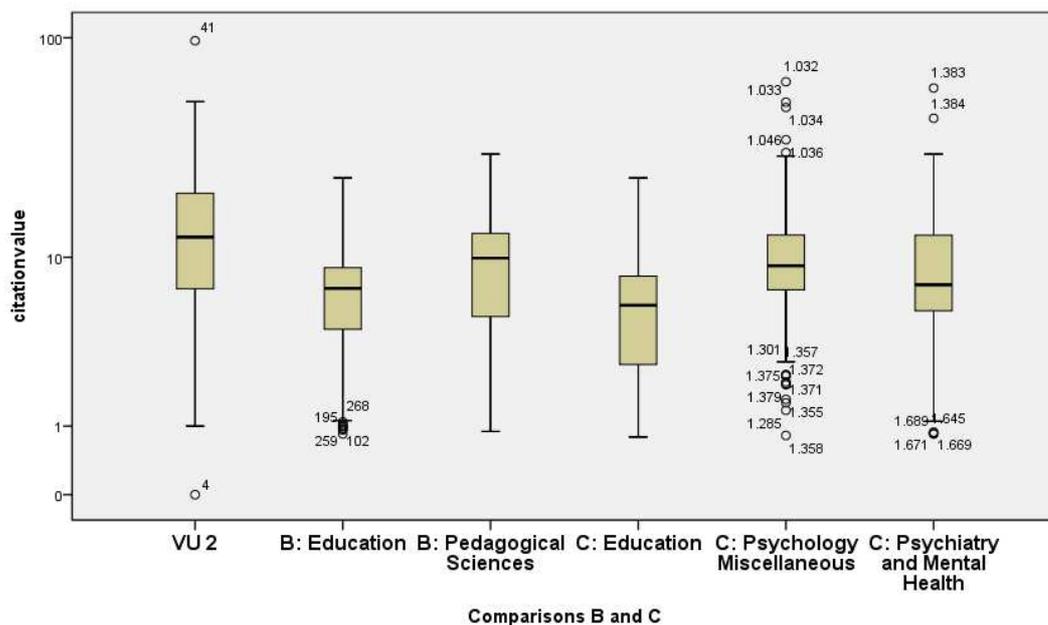


Bibliometric Results VU 2

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
VU 2	28	29	12,4	13	11,5	3	14,81	16
Journals of VU 2	23,21	33,74	18,08	16,43	6,57	2,59	14,36	

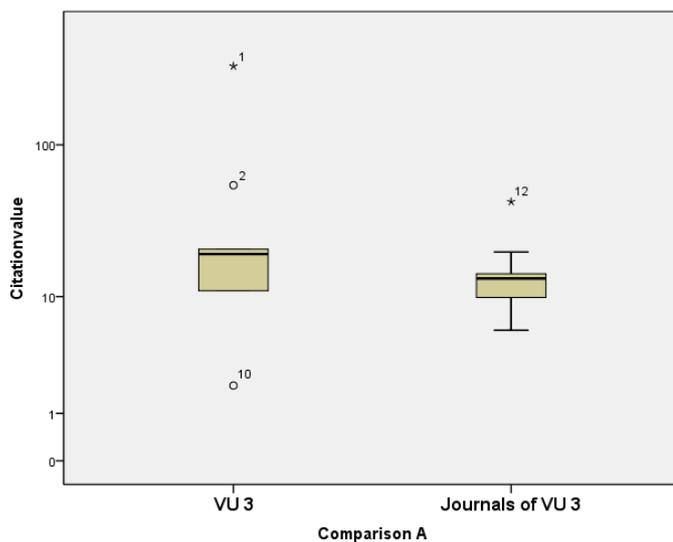


Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
VU 2	27,8	28,7	14,1	10,3	10,8	1,63	16,03	60 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	4%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	21%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	5%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	21%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	20%

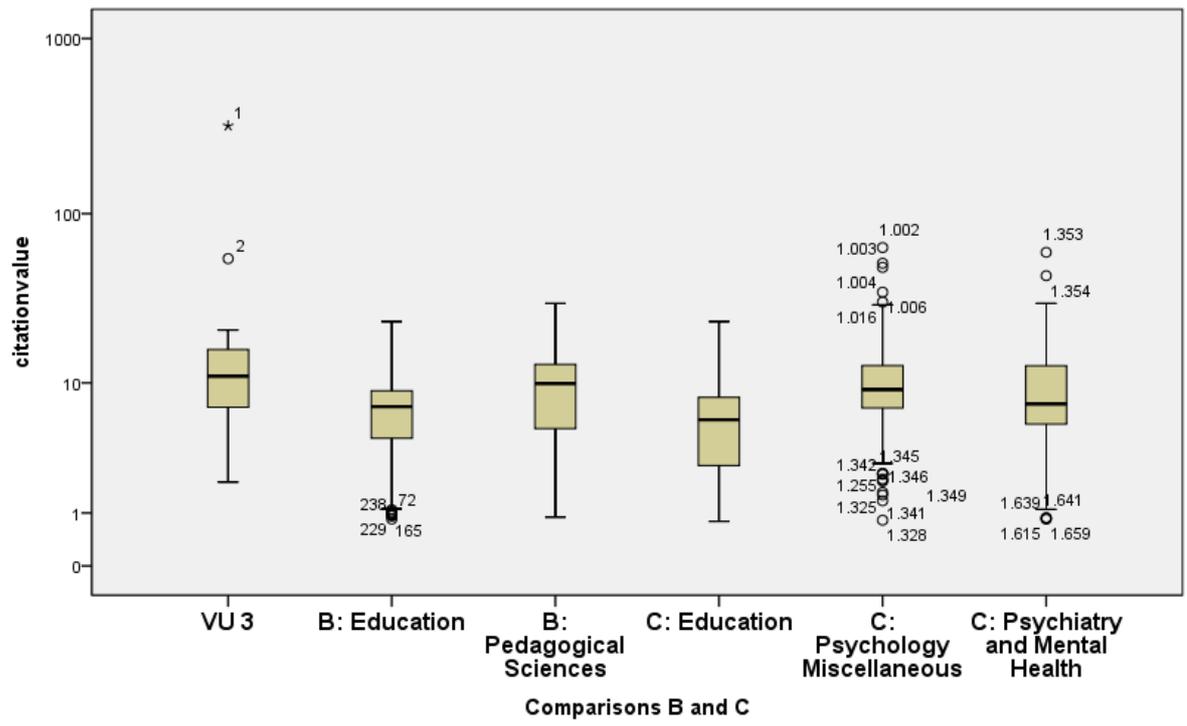


Bibliometric Results VU 3

Comparison A	2006	2007	2008	2009	2010	2011	Overall Average	Articles involved
VU 3				86,6	15	6,5	49,1	10
Journals of VU 3				20,66	12,29	5,73	15,16	



Comparisons B and C	2006	2007	2008	2009	2010	2011	Av. Of Averages	Coverage
VU3				50,6	10,4	7	23,21	29 publications
B: Education	15,88	12,96	9,53	6,44	4,16	1,52	8,42	4%
B: Pedagogical Sciences	22,31	17,62	12,79	8,63	5,56	2,05	11,5	7%
C: Education	15,09	12,08	8,82	5,95	3,7	1,35	7,83	0%
C: Psychology Miscellaneous	22,14	17,73	12,67	8,79	5,58	2,07	11,49	18%
C: Psychiatry and Mental Health	22,75	18,45	12,92	8,54	5,4	2,07	11,69	18%



Addendum 1. Methods

Selection of publications

Each program provided 10 of its most visible publications per year. For all programs but one, the period concerned the years 2006-2011, resulting in 60 publications per program. VU 3 was formed during the period of evaluation, its list spans three years (2009-2011), resulting in 30 publications. In total, 810 publications were sent in. Five publications have been discounted as double entries from the same program. In order to establish that the provided publications indeed relate to the program, publications have been independently checked for proper affiliation of the authors. In a number of doubted cases, program leaders were asked to provide additional information for affiliation of authors. Two publications had to be excluded from further analysis. In total 803 publications have been analyzed.

Some publications featured in the list of more than one program, as these were coauthored by members from different programs. This is the case for 33 publications. Each double entry was from different universities. None were submitted from different programs of the same university.

Scholar as source for impact analysis data

Google Scholar is chosen as the source for data retrieval about the impact of the 803 publications, in order to have a broader view of the impact of the field.³ However, as Scholar results also have been scrutinized for reliability, attention has been given to this issue. Information of all references⁴ to the 803 publications has been downloaded in full for further analysis. To avoid possible errors or bias, Scholar has been used directly, without use of intermediary software such as Publish or Perish. Citations with different spellings of the same publication were included. This was necessary because Scholar ranks its finding according to the results of its indexing. In 14 cases, varying spellings appeared to occur of the same publication. Also, in 27 cases, provided data did not match with Scholar information, usually because of differences in the title. For these cases, search terms for Scholar were corrected if the publication referred to were identical for author, journal and publication year.

In total, 23.905 referring sources were found, of different types. Excluded from the analysis were the references without further detail about their specific (internet)location, i.e. lacking a proper URL, and references predating the year of publication.⁵ This resulted in a total of 21.474 traceable and proper references, 90% of the total number of referring sources.

³ In contrast to Web of Science or Scopus, Scholar includes books, reports, volumes and chapters in volumes as sources for citations, as long as these are made available by publishers and large (university) libraries. For bibliometric analysis of the social sciences Scholar is therefore a more inclusive source of information. Kousha, K., M. Thelwall; Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines, *Scientometrics*, 2008, V73, nr 2, 273-294.

⁴ Terminology: a *reference* is a document or other source listing the publication of which the impact analysis is sought for. This results in a *citation* as the property of this latter publication, i.e. the occurrence of a reference to be found in another source.

⁵ Google Scholar is based on information in books (e.g. Google Books), of libraries and publishing houses. In most cases, this information is traceable and can be considered to represent proper citations. However, as Scholar works as a search engine searching with free format terms, references can also be included that relate to documents that have appeared earlier than the publication of with the citations are searched for.

	Corrected references	Uncorrected references
Article in Scientific Journal	16.473	16.833
BOOK	814	859
CITATION	250	963
DOC	46	121
HTML	762	908
PDF	2.986	4.060
TXT	7	7
Grand Total	21.338	23.751

Table 1. Totals of references per type, corrected and uncorrected

The number of citations to be counted is higher than the above mentioned number of references, since 33 publications were featured by more than one university. The results were 25.233 citations, of which 23.068 citations meet the above mentioned criteria for referencing in Scholar.

For 30 publications no citations could be found in Scholar. In part this may be due to a lack of sufficient indexing by Scholar. As is the case in WoS and Scopus, language bias is also present in Scholar, since indexing by Scholar depends on the cooperation with publishers. This might possibly be the case for 7 of the publications that are in Dutch. Other publications were not traceable likely because these were reports or chapters in volumes not indexed by Scholar. Uncited documents were particularly typical for RUG 1, with 20 publications not to be found cited in Scholar.

One publication received 1460 citations, according to Scholar. However, Scholar gives only the data of the references of 1000 citations. Correcting the total of 1460 references according to the above mentioned criteria was not possible. Therefore, the number of corrected citations for this publication has been manually set to 1314, accounting for the overall 10% overestimation that has been found in other Scholar data.

International Comparisons: Scientific Domains and Domain norms

The goal of this analysis is to provide international comparisons for each of the programs. Since programs are of diverse academic background, it is essential to define the proper context for comparing their impact. In order to provide a baseline for impact performance, three approaches were taken. These were related to the proper context for comparison, the scientific domain in which publications are relevant as scholarly output, and the norm that this context could pose on the investigated publications.

Typically, international comparisons in citation analyses are based on journal impact factors (IF, Web of Science) or journal averages of citations (AC, Scopus) that are characteristic for specific scientific domains.⁶ However, although it is possible to compute such journal averages using Google Scholar for each journal separately, Scholar does not provide information about the averages that are characteristic for scientific domains. Therefore, other sources have been used to arrive at domain specific journal characteristics. Since Scopus data

⁶ See for instance: H.F.Moed, Citation Analysis in Research Evaluation, Springer, 2005.

for scientific domains are freely accessible,⁷ the comparisons in this analysis are based on journal average citations (ACs) derived from Scopus.

Data transformation: Non-linear Multilevel Regression

Scopus measures of impact are different from the Scholar data about journal averages. Scholar data about journals can be computed, defined here as Journal Citation Quotients (JCQ), using Publish or Perish (PoP).⁸ However, the Average Citation (AC) counts in Scopus differ from Scholar JCQ's because ACs are based on averages of citations per documents two or three years, which make these counts moving averages. Scholar JCQ's, by contrast, are cumulative with respect to the number of references because, as a search engine, Scholar doesn't distinguish according to year of publication of the reference. This causes the JCQ for each previous year to be higher than the later year. Using Scopus data as norms for Scholar data requires therefore careful transformation.

In order to perform this transformation, a comparison has been made of both Scholar and Scopus data of a specific set of journals. These were the 72 journals most frequently used by all programs. For each journal a three years Average Citation (AC3) has been obtained from Scimago, based on Scopus data. The three years Average Citation is the total number of citations to each journal in year X divided by the number of citable documents over year X-1, X-2 and X-3. Similarly, using PoP, Scholar journal data have been obtained for this same set of journals. Table 2 provides an overview of these two sets of data of the 72 journals, based on the averages of JCQs and AC3 for these journals (Table 2).

	2006	2007	2008	2009	2010	2011	All year Average
Average of AC3s	2,51	2,49	2,74	2,86	2,96	1,8	2,56
Average of JCQs	34,94	27,01	20,75	13,72	9	3,58	18,17

Table 2. Scopus based three years Average Citations (AC3) and Scholar based Journal Citation Quotients (JCQ), accumulated over 72 journals in Education and Pedagogical Sciences

Non-linear multilevel regression analysis of AC3 of 6 years of 72 journals was used to find a transformation function to bring Scopus AC3 to the scale of JCQ. In this function, AC3 is the Average Citation for a year N, where yrXXXX is 1 for year N, and all other years 0. In $(JCQac3+1)$ correlates 0,876 with $\ln(JCQ+1)$.

$$JCQac3 = \exp(.84 * \ln(AC3+1) + 2.38 * yr2006 + 2.18 * yr2007 + 1.86 * yr2008 + 1.49 * yr2009 + 1.08 * yr2010 + 0.58 * yr2011) - 1.$$

The formula is used to calculate transformed AC3's, yielding JCQac3's that can be used in comparing Group Citation Quotients of programs with data of specific Scopus science categories.

Impact of program

For the comparisons, the impact of a program is computed as the average of the total of citations of the submitted publications of a program, divided by the number of publications involved.

⁷ SCImago. (2007). SJR — SCImago Journal & Country Rank. Data retrieved between September 10 and October 10, 2012, from <http://www.scimagojr.com>

⁸ Harzing, A, Publish or Perish, www.harzing.com.

Addendum 2. Notes on Approaches A, B and C for comparisons

A. *Group Specific Domains*: The first approach is the comparison of impact of programs with the international standards of their own publication world. The diversity of journals and of kinds of output used by programs is large. Of the 125 journal titles in which the programs publish most often, 26 titles are shared by two or more programs, and only nine titles by three or more programs. The basis of the comparison is in this approach formed by the six journals in which each program publishes most often. The average of citations of the articles published in the six most often used journals of each program are compared with the expected JCQ's computed for these journals.

B. *Discipline Specific Domains*: A second approach distinguishes two domains labeled in Dutch "Onderwijskunde" en "Pedagogiek". The names of the domains were translated into English using the terms Education and Pedagogical Sciences respectively. Using Scimago journal listings, journals were selected by researchers of Leiden University from several Scopus categories. These include: Education, Development and Educational Psychology, Clinical Psychology, Psychology Miscellaneous, Applied Psychology and Social Psychology. The resulting 457 journals listed in these categories were subsequently sorted by researchers in the field into the two specified domains. This was done prior to the data collection for group impact. For each year between 2006 and 2011 three years Average Citations of these journals were computed, and averaged for the respective domains for AC3's. Also, for each of the journals in these domains the AC3's per year were transformed into expected JCQ's (JCQac3's), and averages for the respective domains were computed for JCQac3's.

	2006	2007	2008	2009	2010	2011	All year Averages
Average Scopus AC3 for Education	0,7	0,72	0,81	0,86	0,95	0,52	0,76
Transformed AC3 Education (JCQAC3)	15,88	12,96	9,53	6,44	4,16	1,52	8,42
Average Scopus AC3 for Pedagogical Sciences	1,5	1,43	1,5	1,53	1,61	0,91	1,41
Transformed Pedagogical AC3 Sciences (JCQAC3)	22,31	17,62	12,8	8,63	5,56	2,05	11,50

Table 3. Scopus three years Average Citations and transformed AC3 for the domain of Education en Pedagogical Sciences

C. *General Scopus Domains*: The third approach is based on a selection of Scopus science categories that might fit as a norm for the impact of individual programs. Scopus categories are predefined, based on statistics of Scopus Elsevier on a database of over 18.850 journal titles. Relevant Scopus science categories are Education, Psychology Miscellaneous and Psychiatry and Mental Health. Transformation of Scopus data into expected JCQs for each domain is similar to approach B.