

STANDARD EVALUATION PROTOCOL 2009-2015

PROTOCOL FOR RESEARCH ASSESSMENT IN THE NETHERLANDS



STANDARD EVALUATION PROTOCOL 2009-2015

PROTOCOL FOR RESEARCH ASSESSMENT IN THE NETHERLANDS

Published VSNU, KNAW and NWO 2009
(Updated June 2010)

This document is available through internet only: www.knaw.nl/sep

The standard evaluation protocol is developed by the following organisations:

Royal Netherlands Academy of Arts and Sciences (KNAW)

The Academy's mission: As the forum, conscience, and voice of the arts and sciences in the Netherlands, the Academy promotes the quality of scientific and scholarly work and strives to ensure that Dutch scholars and scientists make the best possible contribution to the cultural, social, and economic development of Dutch society. The KNAW's main functions are advising the government on matters related to scientific research, and acting as an umbrella organisation for the institutes primarily engaged in basic and strategic scientific research and disseminating information. www.knaw.nl

Association of Universities in the Netherlands (VSNU)

The VSNU represents the shared interests of the fourteen research universities in the Netherlands in the fields of research, education, knowledge transfer, funding, personnel policy and international affairs. www.vsnu.nl

Netherlands Organisation for Scientific Research (NWO)

The Netherlands Organisation for Scientific Research (NWO) funds thousands of top researchers at universities and institutes and steers the course of Dutch science by means of subsidies and research programmes. NWO comprises eight divisions and acts as the umbrella organisation for nine research institutes. www.nwo.nl

CONTENTS

1. INTRODUCTION.....	3
2. OBJECTIVES OF THE STANDARD EVALUATION PROTOCOL	4
2.1 MAIN OBJECTIVES	4
2.2 CHARACTERISTICS OF THE SYSTEM	4
2.3 META-EVALUATION	6
3. EVALUATION CRITERIA AND FOCUS.....	8
3.1 MAIN CHARACTERISTICS OF THE SEP.....	8
3.2 CRITERIA AND WAYS TO REVIEW THEM	9
3.3 ELABORATION OF THE CRITERIA.....	11
3.4 FIVE POINT SCALE	12
4. PLANNING.....	13
4.1 OVERALL PLANNING.....	13
4.2 PLANNING DOCUMENT.....	13
4.3 SELECTION AND COMPOSITION OF THE EXTERNAL EVALUATION COMMITTEE.....	13
4.4 TERMS OF REFERENCE	14
5. SELF-EVALUATION	15
5.1 CONTENT OF THE SELF-EVALUATION REPORT.....	15
5.2 TABLES TO BE INCLUDED IN THE SELF-EVALUATION	16
5.3 SWOT-ANALYSIS.....	18
6. SITE VISIT AND EVALUATION REPORT	21
6.1 SITE VISIT	21
6.2 EVALUATION REPORT	21
7. FOLLOW-UP	23
7.1 POSITION OF THE BOARD	23
7.2 MAKING THE EVALUATION RESULTS PUBLIC	23
7.3 MID-TERM REVIEW	23
APPENDIX	24
CHECKLIST FOR INTERNAL USE BY THE EVALUATION COMMITTEE	24

1. INTRODUCTION

The Standard Evaluation Protocol 2009-2015 (SEP) is the fourth protocol for evaluation of scientific research in the Netherlands, following the protocols of 1994, 1998 and 2003. The aim of the SEP is to provide common guidelines for the evaluation and improvement of research and research policy, based on expert assessments.

In 2008, an evaluation of the SEP 2003-2009 undertaken by VSNU, KNAW and NWO showed positive results and users emphasized the importance of continuity in the guidelines for research assessment. In this evaluation, universities and KNAW- and NWO-institutes also underlined that the administrative burden should be diminished and that more emphasis should be placed on societal relevance, on positioning and on benchmarking. The SEP 2009-2015 has incorporated these elements. Lessening of the administrative burden is achieved in this protocol by, among other things, limiting the size of the self-evaluation report and by making the midterm review a very light procedure.

The external evaluation of scientific research applies at two levels: the research institute as a whole and its research programmes. Three main tasks of the research institute and its research programmes are to be assessed: the production of results relevant to the scientific community, the production of results relevant to society, and the training of PhD-students. Four main criteria are considered in the assessment: quality, productivity, societal relevance and vitality, and feasibility.

Since the boards of KNAW and NWO and the executive boards of universities are responsible for the external evaluation of the research units under their authority, they are the primary users of the SEP. Regarding the meta-evaluation of the system, the universities are represented in the board of the VSNU. Next to the boards, the research institutes (as units to be evaluated) and the external evaluation committee fulfil key roles in the evaluation process. The SEP therefore provides guidelines for the boards, the institutes and the evaluation committees.

2. OBJECTIVES OF THE STANDARD EVALUATION PROTOCOL

2.1 Main objectives

The SEP 2009-2015 aims at two objectives with regard to the evaluation of research (including PhD training) and research management:

- Improvement of research quality based on an external peer review, including scientific and societal relevance of research, research policy and research management.
- Accountability to the board of the research organisation, and towards funding agencies, government and society at large.

The rhythm of the SEP consists of a self-evaluation and an external review, including a site visit once every six years, and an internal mid-term review in between two external reviews. In the SEP, guidelines regarding assessment criteria, minimum information requirements and the procedure of the external review are formulated.

After the site visit, the evaluation committee will report its findings to the board of the research organisation. The board will publish the report after internal discussion with the assessed research unit and will make its position regarding the evaluation outcomes public. The evaluation report and the position of the board together constitute the results of the evaluation.

Improvement and accountability

The objective of improvement is aimed at both the research and its management. External evaluations are of great value to the institute and its researchers, since international experts in the field formulate recommendations regarding the research, including the strategy and policies which direct and provide the conditions for the conduct of research.

With the external evaluation, the institute and its research groups account for their research activities to the board of the university, KNAW or NWO. In a broader sense, the external evaluations inform funding agencies, government and society at large of the quality and relevance of research activities, thus accounting for the public investments made in scientific research.

2.2 Characteristics of the system

A broad scope

This protocol is primarily directed toward the evaluation of scientific research. Traditionally, such evaluation focuses on the quality of work according to the standards of scientific disciplines and the ways in which results are communicated to a scientific audience. However, 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. With these elements, the evaluation of research has a broad scope.

Furthermore, the conduct of research is by definition an international undertaking. Nearly all fields of research are characterised by an intensive international exchange of ideas and people. The main partners and competitors of the institutes which will be evaluated might be found, not only within the national context, but also in the international arena. The SEP therefore accommodates this broad perspective, requiring institutes to benchmark themselves internationally and requiring boards to convoke international evaluation committees.

Unit of evaluation

The units of evaluation (they will be referred to throughout this protocol as ‘*institutes*’) may vary among the participating organisations – the universities, KNAW and NWO. An institute may be defined as ‘*a group of researchers with an articulated shared mission, operating within one or more research programmes under the same management*’. The board under whose jurisdiction a research institute falls -notably the boards of NWO and KNAW and the executive boards of the universities (they will be referred to throughout this protocol as ‘*board*’) - is responsible for the organisation and proper procedures of the evaluation of that institute. Each ‘*institute*’ will have a director and/or research leader(s) with final responsibility. Throughout this document they will be referred to as ‘*the management*’.

External evaluation committee

A well balanced composition of the external evaluation committee is of the utmost importance. The members of the evaluation committee should be independent of the institute, well acquainted with the current research practice of the discipline(s) and cover the various other areas of the institute’s activities (e.g. PhD training, research in the context of application, provision and maintenance of research facilities for other academic and non-academic target groups). Research management competence is to be represented in the committee. The committee should be able to position the research area(s) of the institute within the international context and should be able to assess the research according to criteria that fit the field’s research practices. The procedure for the selection of the committee is described in *chapter 4*.

Retrospective and prospective

The objectives of accountability and improvement are served by both a retrospective and a prospective evaluation. Both the assessment of past results and recommendations for future research, research management and research policy are of great importance to the research organisation, its researchers and management. Both the retrospective and prospective characters are reflected in the assessment criteria (*see chapter 3*).

Mid-term review

The main objective of the mid-term review is to evaluate the follow-up of the recommendations of the external review committee three years earlier and to formulate future actions. The mid-term review is an internal procedure. The SEP does not present specific guidelines for the mid-term review except that it has to be a light procedure. The board of the research organisation will decide upon the way in which the mid-term review is conducted and which documentation should be provided by the institute. The objective of the mid-term review is further described in *chapter 7*.

Disciplinary and interdisciplinary elements

The SEP is the basis for research evaluation in all research areas, fields or disciplines. It is of great importance that research activities are assessed according to the standards of the specific discipline. The specific character of each field may require emphasis on some elements of the SEP, while other elements may be less relevant to a certain discipline. The fields of languages & culture, humanities & social sciences, natural & life sciences, medicine & health sciences, design & engineering and agriculture & food sciences may each require different approaches to the evaluation. Within these fields, approaches may also vary among sub-disciplines. While the description of the evaluation criteria and information requirements in the SEP are based on the common ground of these disciplines, the institute is invited to present the specific characteristics of the discipline that inform its research and identity in its self-evaluation and facts & figures.

Furthermore, research in the Netherlands and abroad is of an increasingly multi-, inter-, or trans-disciplinary nature. Institutes and research programmes with multi-, inter-, or trans-disciplinary research require special attention in the evaluation. It is, for instance, often more difficult for these groups to show their results through traditional indicators based on publications in high impact journals, and therefore review committees should include evaluators who have a solid experience in assessing such research. The board responsible for evaluating multi-, inter- or trans-disciplinary research should therefore see to adaptations in the standard procedures necessary to assess these particular aspects of an institute's mission, for example with regard to the composition of the evaluation committee or to specific, more qualitative, criteria and indicators.

Specific disciplinary elements regarding an evaluation may be laid down by the board in the terms of reference (TOR) for the evaluation. In the case of the simultaneous evaluation of multiple institutes, the specific elements may be laid down in the *discipline protocol* (see chapter 4).

Research infrastructures

Institutes may have a broader mission than only to conduct excellent research. Some institutes or research groups, for example, have a mission directed at providing and developing various kinds of research infrastructures in the context of the Netherlands, Europe or worldwide. These infrastructures often service both the academic research and R&D in industry or SMEs. Examples are: the research & development of instruments for space research, providing access to historical sources, the creation and maintenance of library collections, providing access to large scale physics infrastructure, design and prototype facilities, and implications for patient care. Quantifying and assessing this kind of output must also be considered while measuring quality, productivity and societal relevance.

2.3 Meta-evaluation

Accountability not only implies obligations with respect to evaluation and publication of the results. It also demands that the research organisations KNAW, NWO and VSNU guarantee that all research institutes within their jurisdiction participate in this systematic cycle of evaluations. Toward this end, three mechanisms for accountability will be in operation:

Schedule of the evaluations planned

At the start of and during the evaluation cycle, each university, KNAW and NWO will provide an overall schedule listing the research units within their jurisdiction and the years of their external evaluations.

Account of the evaluations completed in the annual reports of the organisations

In their annual reports KNAW, NWO and the universities will provide an overview of the external evaluations carried out that year and of the conclusions that the board has drawn.

Review

During its six year cycle, the Standard Evaluation Protocol and the research assessment processes based on the protocol will be evaluated by an independent expert committee. KNAW, NWO and VSNU will organise this review in 2013. In it, the working of the SEP and the impact of the evaluations on the policies of universities and KNAW- and NWO-institutes will be addressed. The results of the review will be made public.

3. EVALUATION CRITERIA AND FOCUS

3.1 Main characteristics of the SEP

The Standard Evaluation Protocol entails three main characteristics:

- **Two levels of assessment:** The assessment takes place at two levels of research organisation, i.e. the level of the institute (or faculty or research school) and the level of research groups or programmes.
- **Three vital tasks:** The assessment regards the three vital tasks of research organisations, i.e. producing results for the academic community, producing results that are relevant for society, and educating and training the next generation of researchers.
- **Four main criteria:** The assessment entails four main criteria, i.e. quality, productivity, relevance, and vitality & feasibility.

At the level of the institute the main evaluation questions regard the management and research policy activities in relation to the mission, strategy and resources of the institute. The four criteria are to be applied in that perspective. A verdict about quality, for example, will take into account how policy measures have contributed to the performance of the institute's members, while a verdict about productivity will entail an assessment of measures that lead to improvement of productivity.

At the level of the research group or programme, the four criteria are primarily applied to the performance of the researchers. The evaluation here will entail an assessment of the output and activities of researchers, both in quantitative and qualitative form, and of the relevance of the work. Issues of policy and management within the programme nonetheless remain important elements of assessment.

In short, the assessment at the institute level primarily focuses on strategy and organisation, whereas the programme assessment primarily focuses on performance and activities of researchers and the results of their work (output and outcome). At the institute level, explicit attention should be paid to the (policy regarding) education and training of PhD students.

Proceeding from the above, the evaluation report will contain two parts:

- **Assessment of the institute level** in terms of the four criteria, with a focus on policy and strategy, preferably in a qualitative manner, identifying the main issues of praise and criticism and putting forward recommendations for improvement. The accent here is on looking forward.
- **Assessment of the research group or programme** according to the four criteria, with a focus on performance, both in terms of scientific achievements and of societal relevance. The committee may use qualitative and quantitative indicators and indications. Each group evaluation ends with a summary in which the four main criteria are translated into a five-point scale: Excellent / Very good / Good / Satisfactory / Unsatisfactory. An extended description of this scale is given in *sec. 3.3*. The committee is requested to consider the full range of the five point scale and apply the criteria according to the descriptions given.

Specific guidelines for the content of the evaluation report are provided in *sec. 6.2*.

As a guideline for both institutes and evaluation committees, the four main criteria are described in the following section, including indications of how to review them.

3.2 Criteria and ways to review them

Criterion 1: Quality

Quality refers to the level of the research conducted by the researchers of an institute and its groups or programmes compared to accepted (international) standards in that field. As a rule, quality is measured by judging the international academic reputation, the position and the output of the unit to be evaluated. However, in case of a national orientation of a research field, the point of reference consists of other groups in the country. When judging research quality, sub-criteria are:

- Quality and scientific relevance of the research
- Leadership of the institute and the individual leadership of the principal investigators, including research policy and research management
- The academic reputation of the researchers
- Organizational aspects of the institute and of the research programmes, such as the human and financial resources
- PhD training in the institute or within research programmes.

Assessment of PhD training

The evaluation committee is requested to focus on the success rates, supervision and organisational embedment of the programme in the research organisation and research activities involved in the PhD-training. Furthermore, the committee is requested to reflect on the availability of educational resources, such as courses and resources for conference attendance. Information on these aspects is to be provided in the self evaluation report of the institute (see 5.1).

The PhD-programmes are to be assessed on the level of the institute. When the institute assessed by the committee is part of one faculty, but the PhD programme is part of broader interfaculty or interuniversity research school or graduate school, the committee is requested to reflect on these links.

In the Netherlands, PhD training is often organized in research schools. These schools can be collaborations of various universities in a particular field or units within a single university (the latter are usually referred to as 'graduate schools'). A research school may seek accreditation once every six years through the ECOS-accreditation (*Erkenningscommissie Onderzoekscholen*), which operates under the auspices of the Royal Netherlands Academy of Arts and Sciences (KNAW). This accreditation focuses primarily on PhD training, but also on the research mission of the research school. Normally, ECOS accreditation is sought after a SEP evaluation. The report(s) of the SEP evaluation(s) may be used as input for the ECOS-accreditation, if not older than three years. When an ECOS-accreditation is sought after a SEP evaluation, it is recommended to include the information requested by the ECOS in the SEP-self-evaluation.

Criterion 2: Productivity

Productivity regards the relationship between input and output. Output should always be judged in relation to the mission and resources of the institute. When looking at productivity in terms of publications of scientific articles and the like, a verdict is usually cast in comparison with international standards of a quantitative nature. However, this is often not possible when looking at other forms of output (for example health protocols, designs, policy reports). Since many institutes will have variegated output and scientific activities, evaluators are asked to also include other forms of (qualitative) information in their as-

essment.

- At the level of the institute, the judgment regards the policy measures to raise the output to the best and most relevant level possible.
- At the level of the research group or programme, both the output directed toward the scientific community and the output for wider audiences are to be judged. Quantitative and qualitative measurements may be used.

Criterion 3: Societal relevance

This criterion covers the social, economic and cultural relevance of the research. Thus, it concerns a great variety of subjects that are both scientifically and socially relevant (global warming, sustainable energy, inequality, governance, migration and integration, quality of life, water, religion, cultural identity, language problems, etc.). In principle, all research activities can be (or become) relevant for these subjects, though this might be more obvious in some cases than in others. When assessing research activities in terms of societal relevance, evaluators are asked to consider one or more of the following three aspects. The three are not mutually exclusive and meant as indicative guidelines. The institute specifies in its self-evaluation report on which aspect(s) it would like to be evaluated.

- *Societal quality of the work.* This aspect refers primarily to the policy and efforts of the institute and/or research groups to interact in a productive way with stakeholders in society who are interested in input from scientific research. It may also refer to the contribution of research to important issues and debates in society.
- *Societal impact of the work.* This aspect refers to how research affects specific stakeholders or specific procedures in society (for example protocols, laws and regulations, curricula). This can be measured, for example, via charting behavioural changes of actors or institutions.
- *Valorisation of the work.* This aspect refers to the activities aimed at making research results available and suitable for application in products, processes and services. This includes activities regarding the availability of results and the interaction with public and private organisations, as well as direct contributions such as commercial or non-profit use of research results and expertise.

At the level of the institute, this criterion is assessed by reviewing the policy measures aimed at enhancing societal relevance, and the societal orientation of researchers and their activities. This includes the institute's policy for making the results of research available to other than academic users (knowledge transfer).

At the level of the research group or programme this criterion can be assessed by reviewing the various kinds of output and activities through impact indicators or more qualitative measurements.

For the assessment of societal relevance, evidence may be gathered through stakeholder surveys, stakeholder conferences, various forms of impact analysis (studies of behavioural changes of groups or institutions, concrete benefits for specific stakeholders), case studies, etc. Several methods have been developed for specific areas (the payback method for health research, for example) and new methods are being developed. More information about these developments is available through links at the SEP-website www.knaw.nl/sep and at www.eric-project.nl

Criterion 4: Vitality and feasibility

This dual criterion regards the institute's ability to react adequately to important changes in the environment. It refers to both internal (personnel, research practice) and external (developments in the field, in society) dynamics of the group. In the self-evaluation, this can best be assessed through a SWOT-analysis.

- At the institute level, the ability of the institute to react adequately to important changes may be shown by the process of establishing research themes, personnel policy, subject choices, concentration of research lines, etc.
- At the level of the group or programme, for example, it may be shown by the way in which projects are professionally managed. This regards an assessment of policy decisions as well as an assessment of project management, including cost-benefit analysis.

3.3 Elaboration of the criteria

The four main criteria described above are elaborated in the following table in terms of a number of sub-criteria and further in terms of aspects that may be considered in the evaluation.

table 3.1 Assessment criteria, sub-criteria and aspects to be considered

CRITERIA	SUB-CRITERIA	ASPECTS THAT MAY BE CONSIDERED
Quality	A1. Quality and scientific relevance of the research	Originality of the ideas and the research approach, including technological aspects; Significance of the contribution to the field; Coherence of the programme; Quality of the scientific publications; Quality of other output; Scientific and technological relevance
	A2. Leadership	Leadership of primary individuals; Mission and goals; Strategy and policy
	A3. Academic reputation	(Inter)national position and recognition; Prominence of the programme director and other research staff; Impact and significance of research results in the field
	A4. Resources	Human resources; Funding policies and earning capacity; Relevance of research facilities
	A5 PhD training	Objectives and institutional embedding; Structure of programmes; Supervision; Success rates; Educational resources
Productivity	B1. Productivity strategy	Productivity goals; Publication strategy; Rewards and sanctions
	B2. Productivity	Scientific publications and PhD-theses; Professional publications; Output for wider audiences; Use of research facilities by third parties
Relevance	C Societal relevance	Societal quality; Societal impact; Valorisation
Vitality and feasibility	D1. Strategy	Strategic planning; Investments and collaboration; Research topics planned for the near future and their perspectives; Flexibility and anticipation of expected changes.
	D2. SWOT-analysis	Analysis of the position of institute and programmes; Analysis of strengths and weaknesses
	D3. Robustness and stability	Research facilities; Financial resources; Staff competition; Mobility and attractiveness; Expertise within the institute

3.4 Five point scale

The evaluation committee will report its findings after reviewing the three vital functions of an institute in terms of the four main criteria. Regarding the institute level it should focus on policy and management questions. The verdict is given in qualitative form, though a quantitative figure may be added according to the scale here under. The board should make this clear beforehand in the TOR. For the assessment of the groups or programmes, the verdict should be cast in both qualitative and quantitative terms. In the text, the most important considerations of the committee should be clarified, while the conclusion should be summarized in a single term according to a five point scale, 'excellent' meaning world class research, and 'unsatisfactory' meaning below acceptable standards. The committee is requested to consider the full range of the five point scale and apply the criteria according to the descriptions given.

For disciplines that operate primarily in a national context, such as Dutch language, or Dutch law, the relevance of international competitiveness is transferred to relevance on a national level. For these disciplines, research should receive the qualification 'excellent' when it is regarded the top group in the country.

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 internationally competitive and makes a significant contribution to the field. Research is considered nationally leading.

3. Good

Work is competitive at the national level and makes a valuable contribution in the international field. Research is considered internationally visible.

2. Satisfactory

Work adds to our understanding and is solid, but not exciting. Research is nationally visible.

1. Unsatisfactory

Work is neither solid nor exciting, flawed in the scientific and or technical approach, repetitions of other work, etc.

4. PLANNING

4.1 Overall planning

The governing boards of KNAW, NWO and the universities will provide an overall schedule for all evaluations that are planned under their jurisdiction, either per year or per three or six year period. The board is further responsible for the planning of each individual evaluation and its follow-up, which also includes a mid-term review that takes place about three years after the external evaluation.

4.2 Planning document

At the start of any external evaluation process, the board will provide a planning document. This includes all major steps to be taken from the start of the self-evaluation to the eventual mid-term review. It consists of:

- Arrangements for the self-evaluation report and other documentation, such as the outcome of the previous mid-term review and other relevant evaluation results
- Selection and configuration of the external evaluation committee
- Planning of the site visit
- Publication of the evaluation results
- Arrangements for the follow-up of the evaluation

The unit to be evaluated provides a self-evaluation document, which is to be endorsed by the board. After approval by the board, the self-evaluation is sent to the external evaluation committee no later than 4 weeks before the site visit. In order to avoid unnecessary information load for the external evaluation committee, the SEP recommends a concise format for the self-evaluation report (see *chapter 5* for a description of the contents). The complete set of documentation deemed necessary can be made available to the committee on a secluded website or other digital (multi)media.

4.3 Selection and composition of the external evaluation committee

The board is responsible for the selection of the chair and further configuration of the external evaluation committee. The unit to be evaluated is invited to suggest committee members. The board may also consult third parties to reflect on the impartiality and independence of the committee chair and members. The board will officially install the evaluation committee, after which it will make a public announcement.

The selection procedure for chair and members has to ensure the competence, expertise, impartiality and independence of the evaluation committee as a whole. In order to meet these requirements, the board and the unit to be evaluated will carefully consider the fit between the mission of the unit to be evaluated and the required competencies, disciplinary expertise and professional backgrounds of the chair and the other members. Preferably, the competencies and expertise required are to be written down in a profile, which serves as a guideline for proposing actual candidates.

It is recommended that the board first invites the committee's chair and that the chair is consulted about other possible members of the committee. The board may consult third parties with respect to the profile of the candidates and the committee composition within the national and international scientific community.

Furthermore, the board is responsible for appointing the staff supporting the evaluation process, in particular the executive secretary responsible for putting the procedural aspects of the evaluation into effect.

4.4 Terms of reference

The board will provide the terms of reference (TOR) for the external evaluation committee. In the TOR, the board will explain

- the main objectives of the SEP (improvement of quality and relevance of research and accountability);
- the procedures of the site visit, including a final program;
- the expectations of the board with regard to the evaluation report, including the use of the 5 point scale for the institute level.

The TOR may entail additional information about the unit(s) to be evaluated or specific elements for the committee to focus on in cases, for example, where the institute engages in multi-, inter- or trans-disciplinary research, wishes to stress specific disciplinary elements (for instance in design & engineering), has a particular role in the broader scientific infrastructure (for instance a library function), or a specific societal mission (patient care, policy advise). When multiple institutes within a discipline are evaluated, such an assignment is part of the *discipline protocol*.

5. SELF-EVALUATION

The information to be presented to the external evaluation committee contains two parts, (i) a self-evaluation report, which includes a SWOT analysis (*see* 5.3), and (ii) a full set of quantitative information concerning the input and output of the group during six years prior to the evaluation.

The first part, the self-evaluation report, including the SWOT-analysis, will be comprised of necessary information and reflective analysis, presented in the concise format below. It stipulates a maximum length for the various components of the self-evaluation report. The second part of the documentation should, if possible, be presented to the evaluation committee through a secluded website or other digital (multi)media.

5.1 Content of the self-evaluation report

The self-evaluation report should give a concise picture of the research group's work, ambitions, output and resources as discussed in the following sections. In the table below, guidelines are provided for the content of the document. The content should cover the elements of assessment as listed in table 3.1, providing both facts and information about the research activities and a reflection on these research activities. The suggested total length of the elements in the self-evaluation report is 4-5 pages for the institute and 4-5 pages per programme, resulting in 20-30 pages for an institute of average size.

table 5.1 Contents of the self-evaluation report

1	Objective(s) and research area	<ul style="list-style-type: none"> • Vision, mission and objective(s) of the institute • Research area and programmes
2	Composition	Composition of the research unit to be evaluated, based on two indications: <ul style="list-style-type: none"> • total number of employees in each job category (including contract-PhD candidates) and • overview of the various sources of financing (internal and external)
3	Research environment and embedding	<ul style="list-style-type: none"> • National and international positioning ('soft' benchmarking based on SWOT-analysis), • number and affiliation of guest researchers (internally and externally funded)
4	Quality and scientific relevance	<ul style="list-style-type: none"> • 3-5 most significant results/highlights relevant to the discipline, per group/subgroup • 3-5 key publications per group/subgroup (references; full text may be published on secluded website) • Number of articles in top 10% of publications relevant to the discipline; ditto for top 25% • 3-5 most important books or chapters of books, insofar as applicable
5	Output	<ul style="list-style-type: none"> • Number of publications (<i>see</i> table 5.3) • Number of PhDs (completed and in progress) (<i>see</i> tables 5.5 and 5.6) • Use (number of users) of research facilities (if part of institute's mission)
6	Earning capacity	Acquiring projects and programmes through competitive funds, public and private, national and international...
7	Academic reputation	Most important signs of recognition for research staff (prizes, awards, invitations to address major conferences, conference organisation activities, editorships, membership of academies)

8	Societal relevance: quality, impact and valorisation	Socio-cultural and/or technical or economic quality, impact, valorisation (<i>see clarification in 3.2 p.10 and below</i>)
9	Viability	Viability of the unit to be evaluated, in terms of resource management, available infrastructure and innovative capacity
10	Next generation	Information about PhD training. Should entail a description of both how the training of PhD candidates and employment of post-graduates are organised and includes: <ul style="list-style-type: none"> objectives and outcomes of the PhD-programme (in particular mission of the programme and career destination of PhD-graduates) institutional embedding supervision: rights and obligations of both supervisors and PhD candidates, educational components success rates (see tables 5.5 and 5.6).
11	SWOT-analysis	Procedure and outcomes of the SWOT-analysis (SWOT analysis proper may be published on the secluded website) Conclusions regarding strategy and activities based on the SWOT-analysis
12	Strategy	Based on the SWOT analysis, see also 5.3

Clarification:

Regarding chapter 6, earning capacity, the institute will provide information about the acquired projects and funds. In some cases earning capacity may also be shown through acquired positions in prestigious projects, as expressed, for instance, by a principal investigator role in worldwide projects (PI-role)

Regarding chapter 8 of the self-evaluation report, the societal quality of research, the institute will provide information in terms of the three aspects mentioned in chapter 3. Furthermore, research output can be presented in various ways. For example, in technical disciplines, one may consider using patents or collaboration with industry; in medical/biomedical research, clinical applications or protocols; in the humanities, exhibitions; in the social sciences, contributions to educational innovation. More information about methods can be found on the SEP-website www.knaw.nl/sep and on www.eric-project.nl

5.2 Tables to be included in the self-evaluation

The following tables are to be included in the self-evaluation. Other quantitative information may be provided through a secluded internet site.

table 5.2 Research staff at institutional and programme level

	Year-5	Year-4	Year-3	Year-2	Year-1	Year now
Entire Institute <name institute>						
Tenured staff (1)	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Non-tenured staff (2)	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
PhD-candidates (3)	#	#	#	#	#	#
Total research staff	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Support staff	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Visiting fellows	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Total staff	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Research programme #1 <name programme>						
Tenured staff (1)	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Non-tenured staff (2)	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte

PhD-candidates (3)	#	#	#	#	#	#
Total research staff	# / fte	# / fte	# / fte	# / fte	# / fte	# / fte
Research programme #2 <name programme>						
...						

Note 1: Comparable with WOPI-categories HGL, UHD en UD

Note 2: Comparable with WOPI-category Onderzoeker, including post docs

Note 3: Standard PhD (employed) and Contract PhD's (externally or internally funded but not employed)

table 5.3 Main categories of research output at institutional and programme level

	Year-5	Year-4	Year-3	Year-2	Year-1	Year now
Entire Institute <name institute>						
Refereed articles	#	#	#	#	#	#
Non-refereed articles (1)	#	#	#	#	#	#
Books	#	#	#	#	#	#
Book chapters	#	#	#	#	#	#
PhD-theses	#	#	#	#	#	#
Conference papers	#	#	#	#	#	#
Professional publications (2)	#	#	#	#	#	#
Publications aimed at the general public (3)	#	#	#	#	#	#
Other research output <specify> (4)	#	#	#	#	#	#
Total publications	#	#	#	#	#	#
Research programme #1 <name programme>						
Refereed articles	#	#	#	#	#	#
Non-refereed articles (1)	#	#	#	#	#	#
Books	#	#	#	#	#	#
Book chapters	#	#	#	#	#	#
PhD-theses	#	#	#	#	#	#
Conference papers	#	#	#	#	#	#
Professional publications (2)	#	#	#	#	#	#
Publications aimed at the general public (3)	#	#	#	#	#	#
Other research output <specify> (4)	#	#	#	#	#	#
Total publications	#	#	#	#	#	#
Research programme #2 <name programme>						

Note 1: Articles in journals that are non refereed, yet deemed important for the field

Note 2: Publications aimed at professionals in the public and private sector (professionele publicaties), including patents and annotations (e.g. law).

Note 3: Also known as "populariserende artikelen".

Note 4: Other types of research output, such as abstracts, editorships, inaugural lectures, designs and prototypes (e.g. engineering) and media appearances.

table 5.4 Funding at institutional and programme level

	Year-5	Year-4	Year-3	Year-2	Year-1	Year now
Entire Institute <name institute>						
<i>Funding:(1)</i>						
Direct funding (2)	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
Research grants (3)	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
Contract research (4)	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
Other (5)	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
Total funding	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
<i>Expenditure:</i>						
Personnel costs	€/ %	€/ %	€/ %	€/ %	€/ %	€/ %
Other costs	€/ %	€/ %	€/ %	€/ %	€/ %	€/ %
Total expenditure	€/ %	€/ %	€/ %	€/ %	€/ %	€/ %
Research programmes						
<i>Funding:</i>						
#1 <name programme>	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
#2 <name programme>	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %
#3 <name programme>	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %

.....						
Total Funding	fte / %	fte / %	fte / %	fte / %	fte / %	fte / %

Note 1: Number of fte PhD candidates can be an estimate, based on the number of PhD candidates given in table 5.2.

Note 2: Direct funding by the university / KNAW / NWO

Note 3: Research grants obtained in national and international scientific competition (e.g. grants from NWO, KNAW and European Research Council)

Note 4: Research contracts for specific research projects obtained from external organisations, such as industry, governmental ministries, European Commission and charity organisations

Note 5: Funds that do not fit the other categories

table 5.5 Standard PhD-Candidates (1)

Enrolment			Success rates				Total			
Start-ing year	Enrol-ment (male / female)		Total (male+ female)	Gradu-ated after (≤) 4 years	Gradu-ated after (≤) 5 years	Gradu-ated after (≤) 6 years	Gradu-ated after (≤) 7 years	Total gradu-ated	Not yet finished	Discon-tinued
T-8	#M	#F	#	# / %	# / %	# / %	# / %	# / %	# / %	# / %
T-7	#M	#F	#	# / %	# / %	# / %	# / %	# / %	# / %	# / %
T-6	#M	#F	#	# / %	# / %	# / %	-	# / %	# / %	# / %
T-5	#M	#F	#	# / %	# / %	-	-	# / %	# / %	# / %
T-4	#M	#F	#	# / %	-	-	-	# / %	# / %	# / %

Note 1: Standard PhD-candidate with employee status and conducting research with primary aim/obligation to graduate; (AiO, promovendus)

table 5.6 Contract PhD-candidates (1)

Enrolment			Success rates				Total			
Start-ing year	Enrol-ment (male / female)		Total (male + fe-male)	Gradu-ated after (≤) 4 years	Gradu-ated after (≤) 5 years	Gradu-ated after (≤) 6 years	Gradu-ated after (≤) 7 years	Total gradu-ated	Not yet finished	Discon-tinued
T-8	#M	#F	#	# / %	# / %	# / %	# / %	# / %	# / %	# / %
T-7	#M	#F	#	# / %	# / %	# / %	# / %	# / %	# / %	# / %
T-6	#M	#F	#	# / %	# / %	# / %	-	# / %	# / %	# / %
T-5	#M	#F	#	# / %	# / %	-	-	# / %	# / %	# / %
T-4	#M	#F	#	# / %	-	-	-	# / %	# / %	# / %

Note 1: Contract PhD-candidates without employee status, receiving external funding or university scholarship, conducting research under the authority of the institute with primary aim to graduate; (beurspromovendus)

5.3 SWOT-analysis

One of the main objectives of the evaluation system is the improvement of research and research management. The self-evaluation report should therefore also entail an analysis of the institute's strengths and weaknesses, and give a perspective for the future. This can be done through an analysis of the strengths and weaknesses and the opportunities and threats in the environment, a SWOT-analysis. This analysis should be conducted at the level of the institute. The board may decide that the analysis is also to be conducted regarding each programme.

Positioning and Benchmarking

The SWOT-analysis is first and foremost an instrument for reflection on the current position and future prospects of the institute and its research programme(s). An important goal of the

SWOT-analysis is therefore to benchmark the institute’s position in the (inter)national scientific arena. The institute is requested to reflect on this position in relation to its main external partners / competitors and compare its mission and main activities with these points of reference. The institute is requested to inform the committee about what it sees as reference point(s) nationally and internationally.

Undertaking the SWOT-analysis

In a SWOT-analysis, an institute (or research programme) analyses itself in four dimensions, two internal (strengths and weaknesses) and two external (opportunities and threats). The questions to be answered in a SWOT-analysis are fairly simple and straightforward. The gathering of information may also be relatively simple, for example, through interviews with relevant stakeholders in and outside the organisation. There are also more comprehensive ways, for example, through surveys and other quantitative techniques. An institute is free to choose a method, as long as the analysis is based on evidence that is transparent for the external evaluation committee.

table 5.7 Examples of questions to be answered in SWOT analysis

Strengths	1	What advantages do you have compared to other research groups in your national and/or international environment?
	2	What do other people see as your strong points?
	3	What relevant resources do you have access to?
Weaknesses	1	Which aspects of your work do you see as sub-standard?
	2	Which aspects of your activities could be improved?
	3	What kind of activities should you avoid?
Opportunities	1	What are the interesting trends that you are aware of?
	2	Where or what are the good opportunities facing you?
		Useful opportunities can come from such things as: <ul style="list-style-type: none"> • Changes in technology and markets on both a broad and narrow scale • Changes in government policy related to your field • Changes in social patterns, population profiles, life style changes, etc. • Local Events
Threats	1	What is the ‘competition’ in your area doing better?
	2	Are there significant changes in the requirements for the work in your field?
	3	Do you have a bad financial situation, and does this regard the lump sum or other money streams?
	4	Do you have problems finding, keeping and replacing qualified personnel?

At the intersections of these four dimensions, four main strategic questions arise, as represented in the following matrix:

table 5.8 dimensions of SWOT

	Strengths	Weaknesses
Opportunities	Strategic question: Which opportunities can be exploited through the strengths of the institute well?	Strategic question: Which opportunities may help overcome weaknesses?
Threats	Strategic question: How can the institute use its strengths to reduce its vulnerabilities?	Strategic question: To which threats is the institute particularly vulnerable and how can the institute overcome these weaknesses?

Strategy for the next period

Based on this analysis, the institute draws conclusions about its position in the national and international arena. It also identifies the elements of strategy, organisation and/or research activities which are to be adjusted in order to meet the external opportunities and threats, reflecting the conclusions of the SWOT-analysis.

6. SITE VISIT AND EVALUATION REPORT

6.1 Site visit

As a rule, the evaluation committee visits the institute. The committee chair, the board and the management of the institute will agree on the programme for the visit. The evaluation committee receives all relevant material (the SEP, the self-evaluation document, the terms of reference for the evaluation and the visiting programme) four weeks in advance of their site visit.

The chair may request, possibly after consulting the other committee members, additional information from the institute or the board. The committee will meet in a closed session prior to the site visit, after being formally installed by a representative of the board. In that closed session, the committee decides on their working procedure for the visit and for writing the draft report.

During the visit, the evaluation committee meets with:

- The director/management of the institute;
- The research leaders of the institute;
- A number of tenured and non-tenured staff;
- A number of PhD-students.

6.2 Evaluation report

The main objectives of the assessment by means of the Standard Evaluation Protocol are improvement and accountability. To meet these objectives, the evaluation committee is asked to write a report that is comprehensive and concise at the same time (max. 15 to 20 pages).

Basically, the report should contain an assessment of the institute focusing on the criteria mentioned in *chapter 3*. Furthermore, the evaluation report should reflect on the strengths and weaknesses of the institute as they emerge from the documentation and the discussions during the site visit. Consequently, the report should also indicate opportunities for improvement, possible threats and recommendations for how these can be counteracted.

Secondly, the report should clearly assess the quality, productivity, societal relevance and vitality & feasibility of the groups or programmes that belong to the institute. The evaluation committee is asked to assess both past performance and future prospects of groups or programmes. The group reports may be confined to 1 page per group, including the assessment by means of the 5-point scale. It is important that the reasons for the given qualification are sufficiently explained in the text.

To enhance understanding by a larger public, especially when the report is read independently, the committee is requested to include a brief description of the institute and its research activities in an appendix. The institute will provide suggestions for the brief description.

Content guideline

A guideline for the content of the report to be written by the evaluation committee:

Part 1. A review of the entire institute, containing:

- A brief description of the institute, its vision, mission and objective(s), and its research activities
- A reflection on the quality (academic reputation, quality of the PhD-training, financial and human resources and research facilities, organisation and internal processes, leadership, national and international positioning)
- A reflection on the productivity (publications, output) and productivity policy
- A reflection on the relevance (in research, in society, and with respect to valorisation)
- A reflection on the vitality and feasibility, and vision for the future (based on the positioning and benchmarking, and also the strengths and weaknesses in the SWOT-analysis: strategy for future years, competitive strength, robustness and stability; earning capacity).

Part 2. A review of each research group or programme of the institute, containing:

- A brief description of the programme, its objective(s) and its research activities
- A reflection on the quality (originality of the research, academic significance, programme coherence, publication strategy, prominence of the researchers), of the research & development performed at the institute, and of the research infrastructure; the management of the research programme; financial and human resources)
- A reflection on the productivity of the research group, the research & development activities and the research infrastructure (quantification of the published output; of the R&D results; the occupancy of the research infrastructure offered by the institute and quantification of its use by third parties)
- A reflection on relevance (of the research, of the R&D, and of the research infrastructure – both for the academic world and for society; valorisation)
- A reflection on vitality and feasibility, and vision for the future (of the research plans, flexibility and anticipation of changes to be expected in the near future).

7. FOLLOW-UP

The assessment follow-up consists of three elements: a position of the board regarding the findings and recommendations of the evaluation committee, the publication of the evaluation results and the mid-term review.

7.1 Position of the board

After the evaluation committee has presented its evaluation report to the board, the board and the management of the institute will discuss the committee's findings and recommendations. After this, the board will formulate its position regarding the evaluation outcomes and the implementation of the evaluation committee recommendations.

7.2 Making the evaluation results public

The report of the evaluation committee and the position of the board regarding the outcomes of the evaluation together form the evaluation results. The board will make the evaluation results public and available for anyone on request. Preferably the evaluation results will be made available through the internet. The committee report is to be published after being presented to the board. The position of the board regarding the outcomes of the evaluation is to be published within a year after the presentation of the committee report.

7.3 Mid-term review

The mid-term review takes place roughly three years after an external evaluation and consequently three years before the next external evaluation. It has both a retrospective and prospective nature and is envisaged as a light procedure. The main objective is to review the follow-up of recommendations from the last external evaluation and to formulate future actions. The mid-term review is an internal procedure. The board decides on the precise form of the mid-term review, the documentation to be provided by the institute and the time path.

APPENDIX

Checklist for internal use by the evaluation committee

The evaluation committee may use the following checklists for the assessment of the institute's level and that of the research group or programme. The members are requested to use these lists individually (that is, before the gathering of the committee) for their provisional judgment and to see them mainly as starting points for discussions with the other members during the site visit. The use of this checklist does not in any way imply that the final score is an average of all scores. This score is only given after careful consideration by the entire committee.

The numbers refer to the five point scale explained in chapter 3. 5 = *excellent*, 4 = *very good*, 3 = *good*, 2 = *satisfactory*, 1 = *unsatisfactory*.

Institute level

How do you evaluate the institute with respect to:	5	4	3	2	1
Quality					
A2 Leadership					
A3 Academic reputation					
A4 Organisation					
A5 Resources					
A6 PhD training					
Productivity					
B1 Productivity strategy					
B2 Productivity					
Relevance					
C1 Societal relevance					
Vitality and Feasibility					
D1 Strategy					
D2 SWOT analysis					
D3 Robustness and stability					

Research Group or Programme level

How do you evaluate the research group / programme with respect to:	5	4	3	2	1
Quality					
A1 Quality and scientific relevance of the research					
A2 Leadership					
A3 Academic reputation					
A4 Organisation					
A5 Resources					
Productivity					
B1 Productivity strategy					
B2 Productivity					
Relevance					
C1 Societal relevance					
Vitality and Feasibility					
D1 Strategy					
D3 Robustness and stability					