



## Impact evaluation of programme review at University of Stuttgart (Germany)

Theodor Leiber, Nana Moutafidou & Bertram Welker

To cite this article: Theodor Leiber, Nana Moutafidou & Bertram Welker (2018): Impact evaluation of programme review at University of Stuttgart (Germany), European Journal of Higher Education, DOI: [10.1080/21568235.2018.1474781](https://doi.org/10.1080/21568235.2018.1474781)

To link to this article: <https://doi.org/10.1080/21568235.2018.1474781>



Published online: 30 May 2018.



Submit your article to this journal [↗](#)



Article views: 8



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



# Impact evaluation of programme review at University of Stuttgart (Germany)

Theodor Leiber <sup>a,b</sup>, Nana Moutafidou<sup>c</sup> and Bertram Welker <sup>d</sup>

<sup>a</sup>Evaluation Agency Baden-Wuerttemberg, Mannheim, Germany; <sup>b</sup>Philosophisch-Sozialwissenschaftliche Fakultät, Universität Augsburg, Augsburg, Germany; <sup>c</sup>Stabsstelle Qualitätsentwicklung, Universität Stuttgart, Stuttgart, Germany; <sup>d</sup>Freelance trainer, Berlin, Germany

## ABSTRACT

The article reports on a case study of impact evaluation of internal quality assurance, i.e. Programme Reviews, which was carried out by the Evaluation Agency Baden-Wuerttemberg and the Staff Unit of Quality Development at University of Stuttgart (Germany) in the form of a methodological before-after comparison. The impact evaluation exhibits that the external evaluations of Programme Reviews were in tune with the idea of a 'critical friend' approach, stressing the positive aspects while including some constructive criticism. It is shown that the evaluations initiated intended constructive discussions about the programmes, while the integration of internal and external evaluations in the Programme Review needs to be improved. A further related result is that programmes need more extensive feedback to improve acceptance from critical corners. Finally, a series of open answers gives indications for possible improvement of the university's quality assurance and organizational structures, including a better and fair integration of students in quality assurance and a more strategic orientation/integration of teaching and learning.

## ARTICLE HISTORY

Received 15 January 2018  
Accepted 7 May 2018

## KEYWORDS

Higher education; impact evaluation; internal quality assurance; programme review; Stuttgart evaluation model

## Introduction

Quality assurance (QA) including quality development are facts of life at German universities. Although not always popular they are central in the areas of teaching and learning. They are part of the accreditation process which largely replaced ministerial procedures to allow programmes to be implemented or not and intend to encourage continuous reflection of study programmes and curricula.

The German system of legally mandatory QA in the area of learning and teaching evolves around the Accreditation Council ('Akkreditierungsrat'), a state-appointed quasi-autonomous non-governmental organization (QUANGO) which in turn certifies individual QA agencies to the creation of assessments reports as a basis of accreditations – either of individual programmes ('programme accreditation') or – as in the case of the

University of Stuttgart – of institutional quality management systems in learning and teaching ('system accreditation') (AC 2013).

The University of Stuttgart received its system accreditation in 2012, thus gaining a high degree of institutional autonomy but also of responsibility vis-à-vis its students who expect excellent teaching and well-organized study programmes. Therefore, the university has to ensure that the process of quality development is lean, effective and seen to produce results – especially the latter claim has been investigated as part of a European project which implemented an impact evaluation of QA that is mainly based on a before-after comparison (IMPALA 2016; ICP 2016). This approach was applied by four partner groups of higher education institutions (HEIs) and QA agencies from Finland, Germany, Romania and Spain, among them the University of Stuttgart and the Evaluation Agency Baden-Wuerttemberg (both Germany).

The present article reflects mainly on the empirical results of impact evaluation of Programme Review at University of Stuttgart which was carried out in the context of the above-mentioned project, while an overview of the methodological core issues of that project (e.g. scheme of before-after comparison approach to impact evaluation; key guidance issues for impact evaluation; description of available data types; contextualization of the attribution problem) is given in another paper of this special issue (Leiber 2018). The article is organized as follows: firstly, the (internal) QA framework of University of Stuttgart, the Stuttgart Evaluation Model (SEM) is characterized; secondly, a description of the methodological implementation of impact evaluation in the present case study is given; thirdly, results of the impact evaluation are presented. The article closes with a summary and outlook.

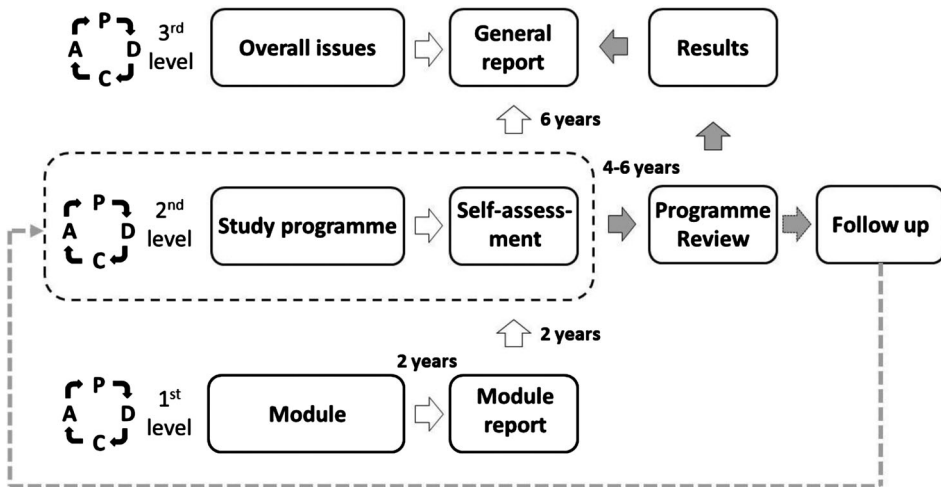
## Quality assurance at University of Stuttgart

### *The Stuttgart evaluation model*

The Stuttgart Evaluation Model (SEM) combines instruments on three levels to ensure QA at University of Stuttgart and has been institutionally accredited in 2012 with a validity period of seven years. SEM works on the three levels of courses and modules, study programmes and the overall performance in learning and teaching on institutional levels (institutes; faculties) (Thumser-Dauth et al. 2013). On each level, monitoring of quality development is established and quality reports are generated on a regular basis, which are part of iteratively applied Deming or pdca (plan-do-check-act) cycles for quality development; these quality reports are also aggregated and fed into the next 'higher' level (Figure 1).

For the present impact evaluation, the focus of University of Stuttgart was chosen to be on the 2nd SEM level, i.e. the Programme Review – a central cornerstone of the quality management system. It replaces (German) programme accreditation by feeding evaluations by external experts into the internal system of quality development.

To see that this instrument is successfully implemented the view has to go beyond the simple compliance with adherence to procedure; ensuring transparency and acceptance of results is central to achieve the assumed positive impact of external elements in QA. Otherwise it will be (or remain) what stakeholders regard as a needlessly time-consuming and meaningless bureaucratic process.



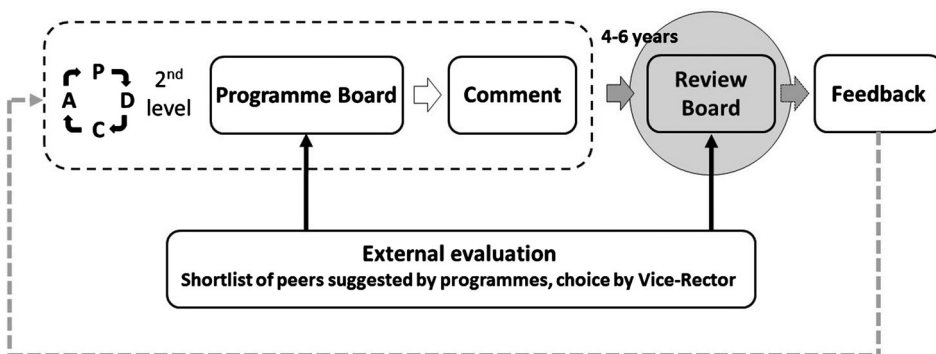
**Figure 1.** The Stuttgart Evaluation Model (SEM) (pdca – plan-do-check-act cycles for quality development).

At University of Stuttgart, the development, implementation and organization of instruments lie with the Staff Unit for Quality Development. Its task is also to gauge the effectiveness of measures, especially given that all SEM processes have been recently established and are ‘work in progress’. Thus, the present impact evaluation was an opportunity to systematically look at acceptance and impact of the first full rounds of Programme Reviews in the SEM.

### *Programme review at a glance*

The Programme Review has been developed to combine internal and external elements, thus replacing the elements of a programme accreditation. It is integrated into a bi-annual cycle of self-assessments which all study programmes have to conduct as part of the system and which takes place every four to six years (Figure 2).

The Programme Review reflects two recommendations made during the system accreditation, i.e. the accreditation of SEM by an accredited QA agency, namely that ‘the



**Figure 2.** The programme review process in the Stuttgart Evaluation Model (SEM).

evaluation of study programmes can be expanded to involve all relevant actors' (Rec. 4) and '[a]s part of a general strategy of curriculum development external experts can be involved on a regular basis' (Rec. 13). Both recommendations are implemented in the review process.

The process itself is based on the self-assessment of the programme and an external evaluation by an academic expert – a 'critical friend' (see, e.g. Balthasar 2011) chosen by university management from a shortlist proposed by the programme representatives. In addition, the Staff Unit for Quality Development provides a check on the formal aspects of the programme and combines all information allowing Programme Boards to comment on the external as well as internal feedback. In a second step all information is discussed in a Review Board chaired by the Vice-Rector. It consists of members from the Senate Committee on Teaching (two professors, one academic staff, one student) and two student representatives from the reviewed or closely related programmes. As part of adapting the process based on various feedback (including the endline survey of the before-after comparison described below) the second round, i.e. Group 2 of Programme Reviews (see Table 1) also included programme representatives in the Review Board meeting to directly discuss issues arising from the information provided. Finally, the Review Board gives a summary of results including possible recommendations for further development or, in cases deemed critical, actual measures to be implemented by the programmes.

### ***The review process – assumptions about aims and structural mechanisms***

Combining the internal checking on formal aspects with the external evaluation by a critical friend the design of the SEM review process is based on several assumptions about intended aims and general structural mechanisms which also guided the present impact evaluation:

- The review process (which lasts around a year in total) should initiate both pre-review and post-review discussions regarding quality and further development on the level of programmes, faculties and university management.

**Table 1.** Time schedule of the programme reviews and the impact evaluation surveys at University of Stuttgart.

Impact evaluation events and QA events	Time/duration
Start of impact evaluation project funding	10/2013
Development of methodology and questionnaires	Circa 6 months
Impact evaluation baseline surveys, Group 1	07–08/2014
Programme Review self-evaluation phase (writing self-assessment report), Group 1	11–12/2014
On-site visit of external peers, Group 1	02/2015
Programme Review, Group 1	03–04/2015
Impact evaluation endline surveys, Group 1	05/2015
Impact evaluation baseline survey, Group 2	07–08/2015
Programme Review self-evaluation phase (writing self-assessment report), Group 2	11–12/2015
On-site visit of external peers, Group 2	01/2016
Programme Review, Group 2	01–02/2016
Impact evaluation endline survey, Group 2	03/2016
Presenting first findings of impact evaluation at conference	16–17/06/2016
End of impact evaluation project funding	09/2016

- The shortlisting of possible external experts by the programme representatives should lead to a higher level of acceptance for the resulting evaluations.
- The internal check on procedural and formal aspects of compliance, which is carried out by the Staff Unit for Quality Development, should allow external experts to focus their evaluation on academic aspects.
- The review results should incorporate the external evaluation and the results of the internal discussions of the Review Boards, thus providing a comprehensive basis for further action.
- The self-assessment report as basis for both the external evaluation (critical friend) and the Review Board meeting should lead to a continuous improvement in the quality of reporting.

## Population, time schedule and sample of the case study

As already mentioned, the methodological core of this impact evaluation consisted of a before-after comparison with a longitudinal panel study with several (online) surveys (Leiber 2018). In contrast to the other three before-after comparison case studies presented in this special issue (Bejan et al. 2018; Jurvelin, Kajaste, and Malinen 2018; Leiber, Prades, and Álvarez 2018), the present study had its focus on case specific questionnaires addressed to students and teachers, while the generic survey questionnaires (ICP 2016, 33ff.), which were used in the other three studies, were only applied at the baseline and, concerning most of the closed questions, were received rather unfavourably by the addressed stakeholders. It is also for this reason that in the present case only baseline and endline surveys were applied. The items of the (non-generic) questionnaire specific to the Programme Review within the SEM are given in Tables 3–9 below.

The analysed Programme Reviews comprised 30 Bachelor and Master programmes from the sciences, humanities and engineering which were grouped into two time-shifted clusters, Group 1 (2014–15: four Programme Boards; 13 programmes) and Group 2 (2015–16: eight Programme Boards; 17 programmes) (see Table 1).

The population size involved only representatives in the Programme Review process, and therefore, from the outset, did not amount to large numbers (Table 2). Moreover, given the ‘popularity’ of quality development as a topic in universities as a whole, the response rates remained relatively low also in the present case study and a numerical interpretation was not applied. Nevertheless, the results can be used to discuss the assumptions regarding Programme Reviews by giving indications. It was also planned to carry out supplementing structured interviews (based on the baseline and endline questionnaires), thus providing more of an idea of the impact, lessons learnt and next steps to be taken. Unfortunately, however, these plans could not be carried out within the funded lifetime of the project mainly for contingent reasons of work overload.

## Results and discussion

### Baseline study

Due to the relatively small, though relevant population (Table 2), there is no possibility to differentiate results according to study programmes. Furthermore, a number of

**Table 2.** Population, samples, response rates and sample errors of surveyed stakeholders at University of Stuttgart for the baseline and endline surveys.

	Population (n)	Sample (n)	Response rate	Sample error
<i>Baseline, group 1</i>				
Students	39	4	31%	24%
Teachers		8		
<i>Baseline, group 2</i>				
Students	85	18	54%	10%
Teachers		28		
Total baseline	124	58	43%	9%
<i>Endline, group 1</i>				
Students	36	9	56%	15%
Teachers		11		
<i>Endline, group 2</i>				
Students	85	10	33%	15%
Teachers		18		
Total endline	121	48	45%	11%
Total	245	106	44%	7%

participants are members of multiple Programme Boards (e.g. Civil Engineering, Real Estate Engineering and Management). Members invited to the survey could therefore give multiple answers regarding the programmes.

At the baseline, students and teachers were asked in the non-generic questionnaire how they experienced the handling of the self-evaluation report of their study programme (Tables 3 and 4).

**Table 3.** Students' assessment of the handling of the programme self-evaluation report at the baseline (percentages and absolute numbers of respondents).

How did you experience the handling of the study programme report?	Totally applies	Applies	Applies in part	Rather does not apply	Does not apply	No answer
There were constructive discussions in the Study Commission about the further development of the study programme during the preparation of the study programme report.	13% (2)	31% (5)	13% (3)	6% (1)	13% (2)	19% (3)
The students of the Study Commission were intensively involved in the preparation of the study programme report.	18% (3)	29% (5)	24% (4)	12% (2)	6% (1)	12% (2)
The strengths and weaknesses of the study programme are well described in the study programme report.	6% (1)	53% (9)	12% (2)	6% (1)	6% (1)	18% (3)
Because of the forthcoming study Programme Review the study programme was reflected from a different perspective.	6% (1)	12% (2)	24% (4)	29% (5)	6% (1)	24% (4)
In the context of the preparation of the study programme report the forthcoming study Programme Review was picked out as a central theme.	6% (1)	29% (5)	12% (2)	0% (0)	6% (1)	47% (8)
As a peer I could gain a comprehensive picture about the strengths and weaknesses of the study programme on the basis of the study programme report.	6% (1)	35% (6)	18% (3)	0% (0)	6% (1)	35% (6)
The study Programme Review will contribute to the further development of the study programme.	13% (2)	31% (5)	25% (4)	0% (0)	13% (2)	13% (3)
I am curious what will be written about my study programme in the external peer report.	59% (10)	18% (3)	0% (0)	6% (1)	0% (0)	18% (3)

**Table 4.** Teachers' assessment of the handling of the programme self-evaluation report at the baseline (percentages and absolute numbers of respondents).

How did you experience the handling of the study programme report?	Totally applies	Applies	Applies in part	Rather does not apply	Does not apply	No answer
There were constructive discussions in the Study Commission about the further development of the study programme during the preparation of the study programme report.	30% (10)	24% (8)	30% (10)	6% (2)	0% (0)	9% (3)
The students of the Study Commission were intensively involved in the preparation of the study programme report.	28% (9)	38% (12)	9% (3)	13% (4)	3% (1)	9% (3)
The strengths and weaknesses of the study programme are well described in the study programme report.	30% (10)	52% (17)	6% (2)	3% (1)	0% (0)	9% (3)
Because of the forthcoming study Programme Review the study programme was reflected from a different perspective.	15% (5)	15% (5)	42% (14)	12% (4)	6% (2)	9% (3)
In the context of the preparation of the study programme report the forthcoming study Programme Review was picked out as a central theme.	27% (9)	24% (8)	30% (10)	3% (1)	6% (2)	9% (3)
As a peer I could gain a comprehensive picture about the strengths and weaknesses of the study programme on the basis of the study programme report.	20% (6)	30% (9)	7% (2)	3% (1)	7% (2)	33% (10)
The study Programme Review will contribute to the further development of the study programme.	19% (6)	28% (9)	28% (9)	16% (5)	3% (1)	6% (2)
I am curious what will be written about my study programme in the external peer report.	31% (10)	38% (12)	13% (4)	9% (3)	6% (2)	3% (1)

**Table 5.** Students' and teachers' assessment of the handling of the Study Programme Review results in the Study Commission at the endline (percentages and absolute numbers of respondents).

How did you experience the handling of the Study Programme Review results in the Study Commission?	Applies	Rather applies	Rather does not apply	Does not apply	No answer
The external peer report was discussed in a Study Commission meeting.	75% (33)	11% (5)	5% (2)	2% (1)	7% (3)
The result of the review meeting was discussed in a Study Commission meeting.	61% (27)	14% (6)	9% (4)	7% (3)	9% (4)

**Table 6.** Students' and teachers' assessment of the handling of the external peer review in the Study Commission at the endline (percentages and absolute numbers of respondents).

How did you experience the handling of the external peer review in the Study Commission?	Applies	Rather applies	Rather does not apply	Does not apply	No answer
The external peer report triggered discussions about the study programme.	39% (17)	36% (16)	9% (4)	14% (6)	16% (1)
Because of the external peer report weaknesses of the study programme were identified.	23% (10)	30% (13)	27% (12)	16% (7)	5% (2)
Because of the external peer report improvement needs of the study programme were identified.	25% (11)	34% (15)	23% (10)	16% (7)	16% (1)
The external peer report confirmed strengths of the study programme.	57% (25)	23% (10)	11% (5)	5% (2)	5% (2)
Because of the external peer report it became obvious that the study programme is on the right path.	56% (24)	26% (11)	9% (4)	2% (1)	7% (3)



**Table 7.** Students' and teachers' expectations referring to the recommendations of the Programme Review Commission at the endline (percentages and absolute numbers of respondents).

What do you expect from the recommendations of the Programme Review Commission?	Applies	Rather applies	Rather does not apply	Does not apply	No answer
The recommendations of the Review Commission concerning technical and content-wise aspects will contribute to the further improvement of the study programme.	21% (9)	30% (13)	23% (10)	16% (7)	11% (5)
The recommendations of the Review Commission concerning formal and organizational aspects will contribute to the further improvement of the study programme.	9% (4)	43% (19)	23% (10)	14% (6)	11% (5)
The conclusion of the Review Commission could be well understood by the Study Commission.	36% (16)	25% (11)	9% (4)	18% (8)	11% (5)

**Table 8.** Students' and teachers' assessments of the study Programme Review at the endline (percentages and absolute numbers of respondents).

	Yes	No	No answer
The study Programme Review made new aspects of the programme visible to me.	43% (19)	48% (21)	9% (4)
If 'yes': The Programme Review made strengths visible.	58% (11)	37% (7)	5% (1)
The Programme Review made weaknesses visible.	84% (16)	11% (2)	5% (1)

**Table 9.** Students' and teachers' assessments of the assessment change referring to the study programme at the endline (percentages and absolute numbers of respondents).

	Applies	Rather applies	Rather does not apply	Does not apply	No answer
Through the Programme Review, my assessment of the programme has changed positively.	9% (4)	23% (10)	32% (14)	16% (7)	21% (9)

In summary, a majority of students think that:

- There were constructive discussions in the Study Commission about the further development of the programmes during the preparation of the study programme reports;
- Students were intensively involved in the preparation of the programme reports;
- The strengths and weaknesses of the programmes were well described in the programme reports;
- The forthcoming Programme Reviews were picked out as central themes when preparing the programme reports;

With respect to opening different reflection perspectives, students were more cautious, i.e. a majority thinks that this is partly or rather not the case. Also, the students (at the baseline) had a high expectation concerning the contribution of the Programme Reviews to the further development of the programmes, and a majority of them was curious about what the external peers would write about their programmes in the external reports (Table 3).

As compared to students, an even greater majority of teachers also agrees with the four bullet point in the list above. In addition, most teachers think that the programmes were reflected from a different perspective because of the forthcoming Programme Reviews (Table 4).

The assessment of teachers and students about whether they could gain a comprehensive picture about the strengths and weaknesses of the programmes on the basis of the programme reports are almost the same. Again, the teachers also had a high expectation concerning the contribution of the Programme Reviews to the further development of the programmes, and the majority of them was curious about what the external peers would write about their programmes in the external reports (Table 4). However, in both items teachers are somewhat more critical or reserved as compared to students (Tables 3 and 4). The latter is also true for the items about the students' involvement in the preparation of the programme report and the possible variation of the reflection perspective (Tables 3 and 4).

### **Endline study**

As already mentioned, regarding the small sample and other reasons of expediency of time, University of Stuttgart decided to focus the endline surveys (for the two programme groups) on a condensed questionnaire with a focus on the review process itself rather than renewed use of the baseline questionnaires.

As mentioned, after the baseline surveys and as a reaction to them some elements of the Programme Review changed: the format of the self-assessment was somewhat altered and programme representatives were actively involved in the Review Board session. Furthermore, information and communication was increased to increase transparency. On the one hand, these changes demonstrate that the baseline study rendered some consequences. On the other hand, however, these changes make a strict and immediate longitudinal view at the data difficult (which the small sample size would have done anyway) but it shows that impact evaluation also has to deal with constant change as an everyday reality of QA and development in a university. Again, the data can still be used to look at the assumptions above and identify possible trends and questions for structured interviews.

At the endline, students and teachers were asked the following questions referring to the study Programme Review:

- How did you experience the handling of the Study Programme Review results in the Study Commission?
- How did you experience the handling of the external peer review in the Study Commission?
- What do you expect from the recommendations of the Programme Review Commission?

The corresponding results are documented in Tables 5–7. Obviously, a great majority of students and teachers experienced that the intended discussions of the external peer reports and the results of the review meetings were discussed in the Study Commission (Table 5). Furthermore, they had the impression that the external peer reports triggered discussions about and confirmed strengths of the study programmes, and they think that because of the external peer reports it became obvious that the study programmes

are on the right paths (Table 6). However, stakeholders are a bit more sceptic about whether the external peer reports can help to identify weaknesses and improvement needs of the study programmes (Table 6).

Students' and teachers' assessments with respect to the possible contributions of the Review Commissions' recommendations to the further improvement of study programmes is tentatively positive but spreads out over all possible answers. Stakeholders see the Review Commission a bit less competent with respect to formal and organizational aspects of programmes as compared to programme contents. A solid majority of stakeholders experienced the conclusions of the Review Commission to be easy to understand by the Study Commission, although critics about this point (circa one quarter) are not negligible (Table 7).

Moreover, students and teachers assessed whether the Programme Review made new aspects, particularly strengths and weaknesses, of the programme visible (Table 8), and whether the review positively changed their assessment of the programme (Table 9). The answers show that more or less half of the respondents experienced that new aspects of the programme became visible by the review. Furthermore, among the yes-sayers a solid majority experienced that strengths became visible, while a large majority experienced that programme weaknesses became visible which underscores the critical, improvement-oriented potential of the Programme Review.

Finally, respondents' change of assessments of the programmes is somewhat indifferent: some state it (rather) changed to the positive, some say it did not (Table 9).

### ***Expenditure and benefit of quality assurance according to teachers***

At the baseline, according to the survey teachers perceived that expenditure of the Programme Review (and SEM) procedures clearly outweighs its perceived benefit, i.e. a majority of about 70% says expenditure is high or very high, while an equally large majority seems to think that benefit is low or very low (Table 10). This is, somehow, corroborated by individual statements of QA-critical teachers in the open answers (see below).

### ***Suggestions for improvements of quality assurance***

In the online survey, students and teachers were also asked whether they have 'suggestions for improvement of procedures of QA and quality development at their HEI', and whether they have any additional comments. In all surveys, a total of 17 students and teachers made suggestions for the improvement of Programme Review (Table 11), while eight students and teachers gave further comments (Table 12).

Again, this case study, even more than the other three before-after comparison case studies (Bejan et al. 2018; Leiber, Stensaker, and Harvey 2018; Leiber, Prades, and Álvarez 2018), cannot be regarded as (statistically) representative. Therefore, all relevant

**Table 10.** Teachers' perception of expenditure and benefit of QA at the baseline.

	Expenditure		Benefit	
	Very low & low	High & very high	Very low & low	High & very high
Baseline	28% (9)	72% (23)	67% (22)	33% (11)

**Table 11.** Students' and teachers' suggestions for improvement of programme review.

---

**Suggestions for improvement of QA procedures**

---

- 'Already in the Study Commissions or in the preparation of the programme report, at least one external member with QA experience should participate, so that external expertise is there and the people take QA then more seriously.'
- 'Currently, the Programme Review is prepared by the professors. Since the students have nothing to say, correspondingly the external reviewers should at least have the opportunity to learn the opinion of the students.' 'The external reviewers should speak separately with professors and students to get a comprehensive picture. At the very least, students should be able to comment on the report of the Study Commission again and this is then submitted to the reviewers.'
- 'There must be consequences, including monetary ones, if a study course does not want to improve or if the relevant stakeholders do not want to participate in the evaluation.'
- 'The study course report is edited jointly by the Study Commission, but unfortunately the student suggestions for improvement and problems are being smoothed, minimized and glossed over. The survey of the students is therefore the great plus of the review. With the help of the review, students' concerns can be brought to the fore from the outside, which are not heard in the Study Commission. In particular, through concrete agreements between the Vice-Chancellor and the Dean of Studies. An improvement would be to allow all student representatives to review.'
- 'Administrative and documentation costs for commissions should be reduced. The procedure is very formalized and bureaucratic.' 'Reduce and compact the complex reports.'
- 

suggestions for improvements of QA and further comments are taken into account, irrespective of the actual absolute numbers of corresponding respondents. Furthermore, since a systematic development of the suggestions/comments from the baseline to the endline is not apparent, this distinction is not made. The respondents' suggestions and comments

**Table 12.** Students' and teachers' additional comments.

---

**Additional comments**

---

- 'Students usually know it best because they are directly affected. Important: As many semesters as possible at a table. If possible also alumni. This is the case with us, so everyone is very satisfied.'
- 'I found the module comments as surprisingly productive and helpful for the further design of the courses.'
- 'Overall, I find it already on a very good level and helpful. In particular, the external expert opinions also bring new aspects (or even things that we do not communicate well).'
- 'At the moment, as a student, you have the impression that you are hardly involved in QA and quality development. You get the evaluation sheet for each event, but you never hear anything from it and generally will not recognize any changes afterwards.'
- 'The procedure is fine. A good Study Commission knows the shortcomings of a course of study. The review does not reveal unknown shortcomings. But the Rectorate is, if necessary, drawn to inadequacies. The Study Commission often lacks the possibilities to remedy the deficiencies. This is also not improved by a review. Nobody asks the Study Commissions/Study Deans how to correct the deficiencies. It is always asked what has been improved.'
- 'In general, a constant rethinking of strengths and weaknesses is very important. This process should be continued so that one does not remain in the same trot.'
- 'Perhaps it would be nice to develop goals from a suitable body (Senate Committee on Teaching and Further Education?), which are to be applied globally for all courses and are looked at in the review. In addition, the subjects must, of course, continue to be aware of the specific objectives. But I would find it very helpful to have global goals and possibly also a prioritization.'
- 'Coaching offers and additional time-limited resources could help the teachers improve the quality of teaching.'
- 'The Staff Unit Quality Development should provide a guide with positioning the objectives of teaching at our university (compared to other universities). The assessment and implementation of these concrete goals can lead the Study Commission and the faculty significantly more efficiently than expensive purchased external assessments without insight.'
- 'One should consider what purpose is pursued. My students find a well-paid job after their studies. And this without quality assurance!'
- 'The whole process is totally overloaded and bloated and has unduly burdened the Study Commission, especially the study managers.'
- 'Evaluate less! Too much effort for little output.'
- 'The Staff Unit Quality Development is increasingly developing into an opaque and over-bureaucratic apparatus that requires a great deal of attention and hardly brings any improvements.'
- 'Let be. It does not matter at all. One of these typical fashion appearances because one is not man (woman) enough to see what is really good and bad. This survey is a typical example: the majority of the questions are concerned with QA itself, with the aim that it is rated very favourably. The questions about the course itself are not to be surpassed in banality. Summary: Waste of time!'
-

were translated from German into English; translation is as close as necessary to the original but with some freedom of abridgement where adequate.

The stakeholders' statements represent opinions of single university members. Taking into account seriousness of the various statements, however, they certainly represent useful indications to check some possible aspects for improvement of Programme Review, such as, e.g. more representative and fair involvement of students; improvements in communication between various stakeholders; (improved) implementation and communication of decisions regarding follow-up requirements (recommendations/restrictions); integration of learning and teaching into comprehensive strategies (Tables 11 and 12).

## Summary and outlook

Going back to the assumptions made above, the following tentative conclusions can be drawn from both the endline studies and a look at the material used (evaluations, self-assessments) regarding the external elements: The external evaluations of the Programme Reviews were in tune with the idea of a critical friend approach, stressing the positive aspects while including some constructive criticism. Both groups, students and teachers, overwhelmingly reported that the evaluations were discussed in the Programme Boards and a majority saw it initiating discussion about the programmes. Both groups again overwhelmingly saw the evaluations as showing the programmes to be 'on the right track'. A very large majority (around 80%) of students and teachers also confirmed that weaknesses were made visible by the Programme Reviews.

Furthermore, by assessment of the Staff Unit of Quality Development, the organizer and moderator of SEM and the Programme Review, four aspects can be concluded concerning effects of Programme Review: Firstly, in general the quality of reporting did increase during the reviews, and the self-assessment reports as basis for both the external evaluations and the Review Board meeting contributed to this – given 'optimal' framework conditions. In particular, according to the reactions of involved programme stakeholders the shortlisting of possible external experts by the programmes led to a higher level of acceptance for the resulting evaluations. In some cases, however, the Programme Reviews also caused a trend of over-reporting (e.g. self-evaluation reports were long and detailed including also content dispensable for development-oriented Programme Review), while generally weak reports did not improve substantially during the process. Secondly, the internal check by the Staff Unit of Quality Development on procedural and formal aspects of compliance allowed external experts (the critical friends) to focus their evaluation on academic aspects. Thirdly, the involvement of stakeholders did improve but there have been problems to ensure adequate student representation and to name professorial members from at least roughly comparable areas for the Review Boards. Fourthly, the integration of internal and external evaluations in the Review Commission's feedback needs to be improved. Programme representatives need more extensive feedback to improve acceptance from critical corners (e.g. programme representatives used the lack of reference to the external evaluation to repudiate the overall feedback). This includes an adequate appreciation of the positive aspects from the external evaluation and a well-argued justification for points raised and recommendations or measures proposed by the Review Boards; therefore, ideally at least one professorial member of the

Review Board should come from a somewhat related discipline. Otherwise acceptance, while always hard to ensure in critical quality cases, will even be less likely to be achieved.

In sum, one could argue that the first two rounds of Programme Reviews at University of Stuttgart show that the review process did indeed initiate discussion, though not always on the quality of learning and teaching. In the case of the second review round, the intention to establish a lean and efficient process showed that this can have several unintended effects. Even with close cooperation and the involvement of programme representatives a lack to transparency in the communication with the Programme Boards might reflect badly on the process itself. Also, the question emerges if the instruments and processes are really accepted if ‘the going gets tough’ and programmes are critically evaluated by the Review Board. Furthermore, the experience of the Programme Reviews, particularly critical cases, illustrates that the organizational issues are likely seen as a more or less harmless nuisance while intervention in questions of content basically calls the expertise of professors and teaching staff into question – a point which is very hard to accept in any expert culture.

In general, as expected, programme representatives (developers, implementers, and managers) will usually always have a hard time taking on criticism and the impact of the process as well as procedural changes should be continuously monitored – a way for which the present impact evaluation has provided an adequate first step. However, we would of course like to use structured interviews to further confirm in more detail that the observed changes were caused by the changes in process, the involvement of programme representatives in the Review Board sessions and the increased communication and support by the Staff Unit.

## Acknowledgement

The authors would like to thank the involved programme commission members of the University of Stuttgart for their willingness and kindness to participate in the impact evaluation surveys and share information. The authors also very much appreciate the support of present and former members of the Staff Unit of Quality Development, Katrin Thumser-Dauth, Susanne Klöpping, Bernhard Minke, and Anne Töpfer. This publication reflects the views only of the authors and the European Commission cannot be held responsible for any use that may be made of the information contained therein.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The authors did part of the work on this article in the context of a project on impact evaluation of quality assurance in higher education institutions, which was co-funded by the European Commission [grant number 539481-LLP-1-2013-1-DE-ERASMUS-EIGF] between 2013 and 2016; Education, Audiovisual and Culture Executive Agency.

## Notes on contributors

*Theodor Leiber* is Scientific Advisor and higher education researcher with Evaluation Agency Baden-Wuerttemberg, Mannheim (Germany) and Associate Professor of Philosophy at University

of Augsburg (Germany). He earned doctoral degrees in Theoretical Physics and Philosophy. His research focuses on evaluation, impact studies, performance measurement and governance, quality management and organizational development in higher education as well as philosophy of science and ethics.

**Nana Moutafidou** is a Scientific Consultant in the Rector's Staff Unit for Quality Development at the University of Stuttgart. Her experiences from several administration units at various universities are important for the establishment of quality assurance related procedures at the university. She was mainly involved in the development and implementation of the Programme Review process.

**Bertram Welker** is a political scientist (Newcastle, UK) and experienced academic administrator. He is currently taking a break from academe and works as a freelance trainer and volunteer.

## ORCID

Theodor Leiber  <http://orcid.org/0000-0002-7218-2603>

Bertram Welker  <http://orcid.org/0000-0002-6975-7779>

## References

- AC (Accreditation Council). 2013. *Regeln für die Akkreditierung von Studiengängen und für die Systemakkreditierung* [Rules for Accreditation of Study Programmes and for System Accreditation]. Drs. AR 2013. Bonn: Akkreditierungsrat. Accessed 24 April, 2018. [http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/AR/Beschluesse/AR\\_Regeln\\_Studiengaenge\\_aktuell.pdf](http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/AR/Beschluesse/AR_Regeln_Studiengaenge_aktuell.pdf).
- Balthasar, Andreas. 2011. "Critical Friend Approach: Policy Evaluation between Methodological Soundness, Practical Relevance, and Transparency of the Evaluation Process." *German Policy Studies* 7 (3): 187–231.
- Bejan, Andrei-Stelian, Radu Mircea Damian, Theodor Leiber, Iohan Neuner, Lidia Niculita, and Radu Vacareanu. 2018. "Impact Evaluation of Institutional Evaluation and Programme Accreditation at Technical University of Civil Engineering Bucharest (Romania)." *European Journal of Higher Education* 8 (3). doi:10.1080/21568235.2018.1474780.
- ICP (IMPALA Consortium Partners). 2016. *Impact Evaluation of Quality Assurance in Higher Education. A Manual*. Accessed 24 April, 2018. [https://www.evalag.de/fileadmin/dateien/pdf/forschung\\_international/impala/impala\\_manual\\_161212\\_v4.pdf](https://www.evalag.de/fileadmin/dateien/pdf/forschung_international/impala/impala_manual_161212_v4.pdf).
- IMPALA. 2016. *Impact Analysis of External Quality Assurance Processes of Higher Education Institutions*. Accessed 24 April, 2018. <https://www.evalag.de/en/international/impact-analysis/the-project/>.
- Jurvelin, Jouni, Matti Kajaste, and Heikki Malinen. 2018. "Impact Evaluation of EUR-ACE Programme Accreditation at Jyväskylä University of Applied Sciences (Finland)." *European Journal of Higher Education* 8 (3). doi:10.1080/21568235.2018.1474779.
- Leiber, Theodor. 2018. "Impact Evaluation of Quality Management in Higher Education: a Contribution to Sustainable Quality Development in Knowledge Societies." *European Journal of Higher Education* 8 (3). doi:10.1080/21568235.2018.1474775.
- Leiber, Theodor, Anna Prades, and María Paz Álvarez. 2018. "Impact Evaluation of Programme Accreditation at Autonomous University of Barcelona (Spain)." *European Journal of Higher Education* 8 (3). doi:10.1080/21568235.2018.1474778.
- Leiber, Theodor, Bjørn Stensaker, and Lee Harvey. 2018. "Bridging Theory and Practice of Impact Evaluation of Quality Management in Higher Education Institutions: a SWOT Analysis." *European Journal of Higher Education* 8 (3). doi:10.1080/21568235.2018.1474782.
- Thumser-Dauth, Katrin, Bernhard Minke, Nana Moutafidou-Heining, and Anne Töpfer. 2013. "Der Studiengangsbericht als Instrument zur Qualitätsentwicklung von Studiengängen [The Programme Report as a Tool for the Quality Development of Study Programmes]." *Zeitschrift für Hochschulentwicklung* 8 (2): 157–166.