Author’s response to reviews

Title: Investigating teaching performance in seminars; a questionnaire study with a multi-level approach.

Authors:

Annemarie Spruijt (a.spruijt@uu.nl)
Jimmie Leppink (jimmie.leppink@maastrichtuniversity.nl)
Ineke H.A.P. Wolfhagen (i.wolfhagen@maastrichtuniversity.nl)
Albert J.J.A. Scherpbier (a.scherpbier@maastrichtuniversity.nl)
Peter van Beukelen (p.vanbeukelen@uu.nl)
Debbie A.D.C. Jaarsma (d.jaarsma@amc.uva.nl)

Version: 2 Date: 19 August 2014

Author’s response to reviews: see over
Dear Dr Aldrin Ulep,

On behalf of all authors, I would like to thank you and the reviewers for your constructive feedback on our manuscript entitled: “Explaining teaching performance in seminars; a questionnaire study with a multi-level approach” (MS: 7214805321285205). We are pleased that the Board has given us the chance to submit a revised version of this manuscript.

We have revised the manuscript in accordance with the comments and requests of the reviewers. In this cover letter, we explain in detail how we have addressed each comment.

We are looking forward to your response.

Yours sincerely,
Annemarie Spruijt, DVM

Changes to the manuscript are highlighted using red coloured text. We have addressed the requests and comments as follows:

**Editorial Comment**

This is a very interesting and useful research work, as recognized by the reviewers. The results achieved are interesting and contribute to the clarification of how contextual elements in seminar groups may condition student perception of teacher performance. However, important revisions are necessary to consider publication. I believe the combination of the 3 reviews provide an excellent critique of the paper. What I would call the main conceptual reasons for concern are:

1) that, as pointed out by Dr Loerstcher, the work is not about observation of teachers performing, but more on ratings of performance made by students? concerning this issue, the paper is too misleading;

Throughout the manuscript we now explicitly state that this study is about students’ perceptions of teaching performance to avoid any misinterpretation. Changes on this subject are made in the Abstract, Background, Method and Discussion section. We for example added the following sentence to the ‘Background’ section:

“In this study teaching performance is measured through questionnaires that are filled in by students right after a seminar facilitated by this teacher”.

2) I have also concerns that the we are not able to assess the quality of the instrument? I suggest that when the authors resubmit a revision, they also submit the paper that describes and validates the instrument;

We have added extra information in the article about the method and results of validation of the instrument so that the quality of the instrument can be assessed better. In the Appendix results of the PFA are presented with the extracting four factors, items, factor loadings, eigenvalues and % of variance explained by the factors. Because the validation process is part of an article that is under review at another journal we are not able to submit this paper to BMC Medical Education. If you appreciate it, we would be able to send this article to you confidentially.

3) the methodological issues, picked up by the referees.

Below, we explain in detail how we have addressed the comments on the methodological issues.

Further to the suggestions of the referees, I would only add one minor: I think the use of ‘seminar learning’ in the first section of the abstract is unnecessary and misleading: the work is solely about "teaching performance"(as assessed by students) which is only one piece in the "seminar learning" experience of students.

We understand your feedback and have changed ‘seminar learning’ in ‘seminar teaching’ to prevent misinterpretation.
Report Reviewer 1

Major Compulsory Revisions

Since “Attending the seminar groups is optional”, this might be a very relevant bias in the results. How did the authors control this?

We think that this bias (students attending seminars versus students not attending seminars) is fairly limited, because the attendance of the seminar groups was on average 79%. Therefore we think that our results can be interpreted as representative. In the ‘Subject and procedure’ section we have added the following information:

“There was an average attendance of 79% of the students during the seminars.”

In the ‘discussion’ we have added: “Because the attendance of the seminar groups was on average 79% we think that our results can be interpreted as representative.”

About the other possible bias (students who have attended the seminars, but did not fill in the questionnaire) we describe in the ‘strengths and limitations’ section “that we cannot make further statements about a possible bias in our population because we do not know the perceptions of students who choose not to participate in the study”.

In line 183, authors mention that “multiple responses from the same students who participated in different seminars…”. I’m concerned with this topic because if the exploratory factor analysis (EFA) was performed assuming that there are students “duplication” this will cause some bias in the analysis.

The exploratory factor analysis we performed for the development of the USEME instrument [23] is indeed for this reason, based on one questionnaire per student. Besides that, the factors that we use in this article are not only based on the exploratory factor analysis of abovementioned article, but also on earlier published research [4,10,11]. We have added the following text in the article:

“The previously developed instrument for evaluating seminar learning (USEME) [23] that was based on previous studies on seminar learning [4,10,11] was used to collect information about the different factors that may explain teaching performance.”

We have no relevant information about the statistical procedures used to perform the EFA (extraction method, rotation, number of factors extracted criteria, etc.).

The extra information of the performance of the EFA is in “Spruijt et al. Factors influencing seminar learning and academic achievement, under review” described as:

“We performed principal factor analysis (PFA) with promax (oblique) rotation to take into account correlations between factors (Wetzel, 2012) and to check whether the items grouped together in the questionnaire were congruent with factors that were found to affect seminar learning in earlier studies (Jaarsma et al., 2008; Spruijt et al., 2012, 2013). The number of extracted factors was based on Kaiser’s criterion (retaining factors with associated eigenvalues of one or larger) and inspection of the scree plot. Although the factor analysis was based on pair wise deletion of missing values, list wise deletion and mean imputation of missing values resulted in highly similar factor loadings and correlations between factors. Because PFA cannot handle repeated measurement structures, only the first questionnaire that a student handed in was used for PFA.”.

In the ‘Variables’ section we have added the following text: “The underlying items, factor loadings and eigenvalues of the instrument that emerged after principal factor analysis with promax rotation and based on Kaisers’criterion and scree plot inspection are presented in Appendix 1 as background information”.

It is relevant to provide communalities scores.

We think that factor loadings provide more information than communalities scores because communalities scores do not differentiate between the loadings on multiple factors. We can understand that this information is relevant for the PFA done in our other study, but we do not think it is relevant for this current study.

Please provide response rate. Is it 988/1582=62%? But this might not be true, because there are multiple responses from same students…

We agree that we have been not clear enough in this. We meant that of the possible 1582 questionnaires (total amount of participants of the seminars) that could have been handed in, 988 completed questionnaires were returned. This was done by 410 different students. We cannot make any statements about how many different, individual students have attended all the different seminars
because we only have information of the students that filled in the questionnaire. Some students may not have attended the seminars in their designated group and our design did only include the amount of participants that attended the seminar groups and did not include an exact overview of which individual students were present at the specific seminars.

In the article we have added: “In total 1582 questionnaires could have been handed in” and “a total of 988 (62,5%) valid sets of questionnaires filled in by 410 different students (median number of 2 questionnaires per student, range 1-6) were collected”.

**Based on table 1, can we assume that the regression model was based on 751 students?**

The regression model is based on 751 questionnaires instead of students. In the article we have added: “based on 751 questionnaires” at Table 2.

**Shouldn’t authors consider year 1, 2 and 3 as a relevant variable in the analysis?**

Please provide the cut-off p values to consider results significant.

We agree that there is a large body of literature on students change in learning strategies over time in other learning contexts, but after the previous studies on seminar learning [4,10,11] we did not have a specific argument that the student’s year of study might influence perceived seminar teaching performance. Because also another reviewer gave feedback on this issue we have added this variable to our database and analysed it. We can now conclude that this variable does not significantly contribute to the prediction of the response variable (teaching performance).

The cut-off values we considered were 0.05. We have added it in the ‘Method’ section in the heading ‘Analysis’: “P-values of < 0.05 were considered significant”.

**The authors don’t provide any information about the statistical procedure assumptions.**

We agree that extra information on our multilevel analysis will clarify the statistical procedure assumptions. We have added:

“Multilevel analysis takes interdependence of ratings into account, in the present study for example, students are nested within groups. Multilevel modelling disentangles these dependencies by quantifying the degree to which variance in student ratings is due to differences between individual students or to the higher level construct that is being evaluated (seminars) and within which ratings are nested. The multilevel method also allows the inclusion of explanatory variables (e.g. seminar group characteristics like group size) that may explain differences between students or between teachers. Another attractive feature of multilevel analysis is that the effects of explanatory variables can be estimated not only as a fixed effect (as in multiple regression analysis) but also as a random effect. This means that it is possible to estimate not only the average effect of, for example, student’s preparation on teaching performance but also the degree to which this effect differs across seminars (random slopes). By choosing multilevel regression instead of multiple regression we are able to deal with the violation of the assumption of independence of observations (students nested in groups, multiple responses of a part of the students) and with deviations of the classical assumption of homogeneity of variances (by specifying random effects)”.

**Authors present on table 1 descriptive statistics for the four scales and group size. I detect some problems regarding normality. There are some skewed distribution and with problematic values for kurtosis. This might be related with the sensitivity of some of the scale items (not possible for me to analyses; I would need descriptive results for all the used items).**

We have added the following paragraph to the ‘Result’ section:

“The descriptives in Table 1 are shown to give a conservative indication of normality because statistics like skewness and kurtosis are based on the assumption of independent observations. As explained above, in our study we don’t have independent observations. Unfortunately there are no alternatives for these descriptive statistics for the multilevel analysis method. Because we did not find extreme values or outliers in these descriptives and because of the large size of this study we can interpret that our method seems fairly robust against the deviations of normality that exist in kurtosis”.

In addition, we think it is not illuminating to provide descriptives on items if factor analysis has shown that certain items can be grouped together to factor scores. In our opinion searching for deviations on item level enhances the chance to find deviations that do not have an intrinsic reason.

**Variables in a multilevel model are most frequently grand-mean centered. I think that this transformation would have been a better option.**
The variables that we describe are grand-mean centered. During the analysis we have added an extra step in order to be able to compare the magnitude of the different predictors. We describe this in the fifth paragraph under ‘Results’:

“For interpretation purposes, we standardized all of the variables (i.e., mean of zero and standard deviation of one), including the response variable. An advantage of using standardized variables is that the regression coefficients of different predictive variables can be compared in terms of magnitude. A larger coefficient indicates a stronger contribution to the prediction of the response variable”.

Based on the previous point, I don’t understand the authors’ sentence: “Table 1 indicates that no variables displayed extreme skewness or average scores that were very close to one of the limits of the theoretical range (i.e., 1-5).”

Considering the comments above, we understand that this gives confusion. Because we now explain this in more detail, we have deleted this sentence.

There are some information that should come on analysis section and not on results (from line 209 to line 217).

We have integrated this information in the analysis section (see highlighted red coloured text in manuscript).

Authors don’t provide relevant information about the model fit.

Information about model fit is needed when dealing with latent variables (e.g., Structural Equation Modeling, Confirmatory Factor Analysis, Latent Class Analysis), but not in multilevel regression analysis and other regression models where no latent variables are used.

Minor Essential Revisions

Please replace “eigen values” for eigenvalues, line 158 and Appendix 1.

We have replaced them.

Report Reviewer 2

Major Compulsory Revisions

Authors should change “teacher performance” to “student perception of teacher performance” (or a similar term) in all cases throughout the paper. This change will require rethinking of the manuscript title and may require reworking some of the text in areas where this term is used. The authors should moderate claims about student learning (for which they have no evidence) like that made in the first paragraph of the discussion.

Throughout the manuscript we now explicitly state that this study is about students’ perceptions of teaching performance to avoid any misinterpretation. Changes on this subject are made in the Abstract, Background, Method and Discussion section. We for example added the following sentence to the ‘Background’ section:

“In this study teaching performance is measured through questionnaires that are filled in by students right after a seminar facilitated by this teacher”.

Authors should include a brief paragraph reviewing what is known about how student perception surveys correlate with other measures of teacher performance and student learning. This paragraph could be included in the introduction or discussion.

In the ‘Background’ we have added the following paragraph on this topic:

“In this study teaching performance is measured through questionnaires that are filled in by students right after a seminar facilitated by this teacher. For students see a great deal of teaching, they are in an unrivalled position to comment on its performance and quality. Besides that, noneexperts in a subject are uniquely qualified to judge whether the instruction they receive is helping them to learn [19]. Specific theoretical and empirical basis in the work of Ramsden and Entwistle (1981) [20] and subsequent studies [21,22] have shown associations between the quality of student learning and students’ perceptions of teaching [19]”.

4
The term “explaining” in the title is vague. It should be replaced with “investigating” or a similar term.

We have changed the word “explaining” in “investigating”.

The authors use the term “facilitating” on pages 3, line 61, page 11, line 282, and page 12, line 316. I stumbled every time I read this and would suggest using “facilitation” instead.

We have replaced the term “facilitating method” by “facilitation method”.

The phrase “deep learning approach” is in quotation marks on page 4, but it unclear where this quote comes from.

We have added the reference behind it.

The term “seminars” is italicized on page 4, line 88. There is no need for italics here.

We have changed it in regular text.

The term “dimensions” used on page 5, line 121 seems odd. The term “characteristics” would be a better choice.

We have replaced the word “dimensions” by “characteristics”.

The term “enroll” is misspelled on page 6, line 135.

We have changed it in “enroll”.

The term “previously” would be more appropriate than “earlier” on page 7, line 155.

We have replaced the word “earlier” by “previously”.

Replace “you” with “she or he” on page 11, line 267.

We have replaced “you” by “she or he”.

Replace “of” with “on” on page 11, line 271.

We have replaced “of” by “on”.

The phrase “insights in and between these variables” on page 11, line 273 does not read well.

We have deleted “and between” to make the sentence easier to read.

Replace “to study to what extent” with “investigation of” on page 12, line 318.

We have replaced “to study” by “to investigate.”

Replace “Descriptives” with “Description” in the title of Table 1.

We think that “descriptives” is a correct statistical term and we therefore did not change it.
Report Reviewer 3

**Major Compulsory Revisions**

USEME instrument Under review elsewhere, means that I cannot estimate the quality of the instrument and thus I do not have insights in the quality of the methods and results in this study. I would suggest to send a copy of the paper in review to be able to understand the instrument used in this manuscript.

We have added extra information in the article about the method and results of validation of the instrument so the quality of the instrument can be assessed better. In the Appendix results of the PFA are presented with the extracting four factors, items, factor loadings, eigenvalues and % of variance explained by the factors. Because the validation process is part of an article that is under review at another journal we are not able to submit this paper to BMC Medical Education. If you appreciate it, we would be able to send this article to you confidentially.

**Discretionary Revisions**

*Introduction: Content does not feel right when applied as a noun. I am not an English native, but I prefer ‘topic’ or ‘subject matter’.*

In sentence 78 we have replaced the word “content” by “subject matter”.

*Lines 78-79: …. may prompt students to adopt a deep learning approach. By using the words ‘may’ and ‘adopt’, it feels like students make an active decision to apply deep learning strategies, while it is not only the student that assures engagement. Sometimes, it’s only the format of the seminar or the teacher that prompts students to engage. Therefore, I would prefer the use of ‘facilitate a deep learning approach’ to ignore who or what makes students engage.*

We realise that this sentence causes confusion. We have rewritten it as:

“There is a tendency to facilitate a ‘deep learning approach’ when participating in the group prompts asking questions, engaging in discussions and interacting with the subject matter.”

*Line 96: ‘roles’ : do teachers change their role in the seminar? Or do they need a variety of skills to ensure learning among the students? In the discussion you mention teachers’ skills (instead of roles?).*

We agree that there is a difference between teachers’ roles and teachers’ skills. Teachers need to change their role during seminars en therefore they need to use a variety of skills to be able to ensure learning.

*Line 101: ….to support the facilitation of the group. Isn’t this a tautologism?*

In this sentence we wanted to say that the rationale supports the facilitation of the group. We have changed the sentence in:

“…they lack a clear standard procedure (like the 7-jump in PBL) [14] that supports group facilitation”.

*Line 237-238: ‘in order to get a deeper understanding of seminars’ or seminar teaching?*

We have replaced “seminars” by “seminar teaching”.

*Line 252: ..... different than our expectation. = ‘against our expectations’?*

We have replaced “different than our expectation” by “against our expectations”.

*Line 278-279: We think active group learning methods, as seminars, are good methods for this to exercise.....I would suggest “ We believe that.... To approach the complexity of real-life problems.....”*

We have changed the sentence as following:

“We believe that active group learning methods, as seminars, are good methods to approach the complexity of real-life problems”.
Line 280: …conflicts about knowledge. What do you mean an assignment in conflict with the students’ cognitive knowledge web or an assignment testing the students knowledge or an assignment?

We agree that the word conflict may confuse in this sentence. We have adjusted part of the sentence in:

“… provided the content holds for challenging assignments that are complex and integrate knowledge”.

Line 280: These kind of assignments demands…. I would suggest: These assignments need to demand….

We have changed it in the way the reviewer suggested.

Teaching performance is measured as ‘didactic performance’ and ‘content expertise’. This is in sharp contrast to the large ‘variation’ of teacher’ roles described in the introduction. Can you explain why teaching performance should only be measured by the applied scale?

The reason that teaching performance is only measured by the applied scale lies in the fact that the large variation of roles that are described in the introduction are based on qualitative studies. During the principal factor analysis of the previous study [23] a part of the items on the different roles of the teacher did not reveal as part of the factor ‘teaching performance’ (for example because of inadequate loadings). To ascertain the most valid way to investigate ‘seminar teaching performance’ we therefore used the scale that is developed by a triangulation of qualitative and quantitative methods.

Minor Revisions

Context on teaching performance is rather understudied. However, the context on group interaction is also understudied, as well as the relation between both context on group and teacher & influence of the teacher on the group. This has not included the variable time on all these contextual factors, members of the group and the teacher. Therefore, the representation is a little simplified in the paragraph starting at line 103. Can you add literature and complexity of group learning and the role of the teachers? A figure might help to also explain the need of multilevel analysis at all times.

We agree on the reviewer that these items are understudied and that (research on) group learning is complex because of the many interactions between variables that play a role. In this article we made a deliberate choice to not include the large number of studies on group learning and the strong body of evidence that shows that effective learning in groups is dependent on the quality of interaction, motivation and beliefs of students, quality of tasks and resource, group process and the quality of instruction. We choose to focus on what is known specifically for seminar learning and made the quality of instruction/teaching performance the central variable in this study. To meet the reviewer’s feedback we inserted extra information on the complex interactions between the variables that we know on seminar learning.

“The relationships between these variables are rather complex because of the many interactions between them. For example, in a qualitative study [10] students explained that the opportunity for and quality of group interaction was a combination of the teacher, the size of the group, the motivation of students, group dynamics, amount and type of seminar questions and the extent of preparation of the students. The extent of preparation depends, among others, on the quality of preparation materials provided by the teacher”.

To explain the need for multilevel analysis in studies on group learning we added extra information on multilevel analysis in the ‘Analysis’ part of the ‘Method’ section.

Can you provide some information on the seminar allocation or prescription? For example, do students choose in which seminar they participate? Above all, which seminar teacher they want to go to? And does a student know which students are in the seminar? How is the participation of students in the seminars (% of students who participate and who leaves after participation in the initial session(s))?  

Every semester students are allocated to one permanent seminar group of 25 students by the Office of Educational Affairs. During this semester, which contains five courses, these students participate in multiple seminars with the same student group. However, in another study on seminar learning [10] one of the students said that “Some students do not always attend a seminar in their designated group”. In practice, group composition can thus differ a little. The Office of Educational Affairs makes a time-table for every seminar group. Students do not know who will teach the seminar session in advance. So, in short, students know the other students in the group but do not know the teacher in advance.

We have added some sentences to the ‘Educational context’:

“Every semester students are allocated to one permanent seminar group of 25 students by the Office of Educational Affairs. During
this semester, these students participate in multiple seminars with the same student group, but teachers vary depending on the seminar theme. The Office of Educational Affairs makes a time-table for every seminar group. Students do not know who will teach the seminar session in advance”.

**Line 176-177:** ‘The questionnaires were only analysed if the consent form was signed.’ How many questionnaires were excluded?

We have added the following sentence to the ‘Result’ section:

“2 questionnaires were excluded because they did not have an attached signed consent form”.

**Can you enlighten what it a range of 1-5 means? For example, if 3 is neutral, is it correct that the mean number of students are just sufficiently prepared?**

When scoring the items grouped within the variables/factors displayed in the Table, student rated ‘1’ if they ‘totally disagreed’ on the item and ‘5’ if they ‘totally agreed’. ‘3’ is ‘not disagreed/not agreed’. In the descriptives of the factor scores displayed in Table 1 the average scores of the items belonging to the specific factors are shown. The mean of the factor student preparation is 3.48, but because of the standard deviation of 0.91 (which has to be interpreted with cause because of the non-independent questionnaires – see above) it shows us that not all the 751 students said to be well prepared.

**Can you show how many students have been involved in two or more seminars that were ‘selected’ to participate in this study?**

We added the following sentences to the article:

“A total of 988 (62.5%) valid sets of questionnaires filled in by 410 different student (median number of questionnaires per student was 2, range 1-6) were collected”.

**Have you done any sub analysis between the year groups? There is a large body of literature that students change in their learning ‘strategies’ over time. So I would like to see sub analysis to see if students in year one value different contextual factors important for teaching performance.**

We agree that there is a large body of literature on students change in learning strategies over time in other learning contexts, but after the previous studies on seminar learning [4,10,11] we did not have a specific argument that the student’s year of study might influence perceived seminar teaching performance. Because also another reviewer gave feedback on this issue we decided to nonetheless add this variable to our database and analysed it. We can now conclude that this variable does not significantly contribute to the prediction of the response variable (teaching performance).

**The teacher intercept is relatively strong compared to intra-student and intra-seminar group intercepts. Can you explain or enlighten this in the results section (and discussion)?**

We have inserted the following sentences:

“The larger variance around the random intercept teacher seems to show that there is more variance in that level, but we need to be careful with statements on causality because a part of the random intercept variance can still be explained by other variables that we did not measure in this study. These ‘unmeasured’ variables can be variables on teacher level, but also on seminar or student level”.

**Second paragraph of the discussion – group interaction : I would prefer emphasis on how these results align with the literature. A large body of literature shows how group interaction increases student learning. As seminars aim to increase student learning, group interaction is indeed the best way to reach this goal (largest effect).**

We added some sentences on this topic in the second paragraph of the discussion.

“Since empirical work of Ramsden and Entwistle (1981) [20] and subsequent studies [21,22] have shown associations between students’ perceptions of teaching and the quality of student learning [19] and a large body of studies have shown that group interaction increases the quality of student learning [24,25] this can be another indication that group interaction within seminars also enhances student learning in seminars”.

**The last sentence does not add to the point of discussion in the paragraph. Can you enlighten why it is relevant to emphasize the need to apply group dynamic skills to increase tutor performance?**
We have deleted that last sentence and enlighten the relevance of emphasizing the need to apply group dynamic skills in the second and fifth paragraph of the ‘Discussion’.

Third paragraph – preparation: These results indeed show an ambiguous relation between preparation and repetition of the subject matter by teachers. Although it is not the aim of this study, it would add insights for teachers if you can add a suggestion how to deal with these differences in the seminars.

We have added the following text to the third paragraph:

“One example to prevent these different levels of knowledge might be the introduction of initial prior knowledge assessments before or at the beginning of a seminar (comparable to the Readiness Assurance Test that is use in team-based learning curricula)[27]. Teachers can also decide to approach the less prepared students other than the well prepared student. The teacher can discuss the seminar content with a subgroup of well-prepared students while the unprepared students discuss the content with each other and may only consult the other well prepared group of students before they consult the teacher”.

Paragraph starting at line 262 – group size: The effect size is rather small. This does not return in the discussion. Could you add this in the discussion? The references in this paragraph do not include the variety of literature on group size. Can you include a wider variety of group size research to reflect the current knowledge on group size and student learning. This will also make your inferences stronger why groups need to be relatively small. In my opinion, the question remains unanswered ‘what is the critical group size for seminars’? Closer to 20 or closer to 30 students? Can your results help to understand the critical group size in seminars?

We have added the word “small” and the following wider variety of group size research in this section:

“This result did not surprise us because there is a general consensus that the optimum size for face-to-face small group meetings is between 5-8 [5,29,30]. Besides that, for a teacher leading discussions between students it is easier in a small group because she or he can facilitate active participation and knowledge construction of the students better. However, it is stated that groups can function productively and satisfactorily even if their size lies outside those limits, but the facilitator has to work harder and need to show leadership skills more [5,30]. We therefore think that by dividing the seminar group in subgroups you have can reach ‘best of both ways’.”

With the studies we have done so far, we cannot propose a recommendation for the critical group size for seminars yet. We have added your suggestion to ‘suggestions for future research’.

Line 276: I would exclude the brackets, as ‘questions’ are a good method to engage students, however, knowledge questions or questions only are not sufficient to facilitate deep learning.

We have deleted the brackets.

Paragraph starting from line 270 Recommendations: I like these recommendations, but the sequence of all recommendations is quite long. Can you use some space in between the recommendations to ease the various subject areas in which the recommendations are given? And can you announce the recommendations in the paragraphs above so this paragraph is not a surprise.

We have applied these tips.

In the discussion I miss the discussion on the medium effect of the topic of the seminar on teaching performance. Does this imply that the topic is independent from the method of the teacher? In other words, some topics will always be boring? Or are some topics difficult to make interesting than others?

The reviewer is right that ‘interest in subject matter’ is part of the factor ‘content’ we have studied. However, in this study we chose to investigate seminar teaching in general by a quantitative study. Because of this chosen design we cannot make any statements about individual seminars.