* Manuscript for ICEE2015

Measuring learning gains in project management

S.A. Nijhuis[1], R. Vrijhoef[2] and J.W.M. Kessels[3]

1 Utrecht University of Applied Science, Netherlands, steven.nijhuis@hu.nl

2 Utrecht University of Applied Science, Netherlands, ruben.vrijhoef@hu.nl

3 Twente University, Netherlands, joseph@josephkessels.com

Abstract

Teaching project management is becoming a standard part of curricula in higher education. Assessing the added value of the teaching efforts needs pre- and post assessments. Given the wide variety of skills and knowledge project management embraces a proper assessment of project management is difficult. A method of assessing added value has been designed and tested on the first part of a professional Master in Project and Process Management. The design is based on students assessment of learning gains (SALG) with several extra criteria. The design was evaluated, updated and tested again. The third test with a tweaked design is being performed. The results do not convince that this SALG-based instrument can be used to measure added value.

Keywords: project management, added value, measuring, education, students assessment of learning gains.

1. Introduction

Crawford, Morris, Thomas, and Winter [1] state: ‘Project management is offered as a significant component in a range of undergraduate and postgraduate academic qualifications, including construction, engineering and IT,’ which is in line with the desire to make the higher education studies more relevant to daily work practice [2]. Martin [3] claims that project management is an important element of both management and engineering education. But incorporating project management is not easy, as Ellis, Thorpe and Wood [4] note: ‘Project management is a challenging subject to deliver, not least because of the wide variety of skills and knowledge it embraces.’ Michel and Prévôt [5] note that more emphasis should be placed on soft skills such as social, emotional and organizational.

There is a great variety of suggested ways to convey project management competences: simulation training, service learning, PBL, PBL with a project manager from a different study, case study, etc (see [6] for a more extensive list of literature). Most of these alternatives are not evaluated on the success of the pedagogical approach, but by means of student enthusiasm (‘I learned a lot’) and/or 'ticking off' products (planning, report).

Measuring and evaluating the success of a pedagogical approach /…/ is crucial as it allows one to determine if the given approach is indeed effective, with objective measures to accompany the claim [7]. Although assessment has been implemented for a long time in project management [8], it is found to be a difficult task to perform correctly in higher education [9]. The required behavioral skills

… are difficult to assess and innovative approaches are required [10]. It can be argued that assessing competences of project managers is already available for certification purposes, but those assessments are targeted at project managers with experience and not at student level or are aimed at the knowledge component only. A link between these certification systems and achieving project success is almost nonexistent [11,12].

Measuring learning outcome is not sufficient for evaluating teaching efforts, we need to assess added value to evaluate the teaching efforts, as the incoming ability is the largest predictor of the outcome ability [13].

A relative simple form of a performance test is a 360 degrees feedback system where the student is evaluated by a group of peers. In new groups, where students do not have ample previous experience with each other a 360 degrees feedback system will not be able to produces a pre-test and therefore will not be able to measure learning gains. Moreover peer performance test can suffer from a halo effect [14]. More elaborate performance tests (like role play or expert assessment) have the disadvantage of being labor and/or money intensive [15,16].

In this paper we report on design research [17], aimed at using Students Assessments of Learning Gains (SALG) [18], in order to measure learning gains in a professional Master of Project and Process Management. The design has been tested and evaluated twice and is being performed on the third group of students.

First we introduce the design criteria for using SALG, secondly we describe the case study and thirdly we describe the tests and evaluations. Limitations are discussed before we reach a preliminary conclusion.

2. Assessment of learning gains

Considering that most higher education institutions do not have an infinite funding, a labor intensive assessment for a pre- and posttest will usually not be used to measure learning gains, and this is reflected in literature where various researchers turn to SALG [7,18-22] to illustrate or claim teaching effects. There is a difference in how SALG is used. Most don’t use a pretest, some ask students what activities helped the learning process [20] and some ask students their perceived learning gains in a posttest [19,22]. SALG does have a big weakness, since it is self assessment. But 'though it has some limitations [it] may be used as part of a multi-source evaluation scheme' [23]. They go on to conclude 'Most studies of self-assessment are in areas of technical knowledge and ability. Even in concrete areas such as these, self-assessment has been found to be inaccurate'. Noting this risk, we will develop design criteria for using SALG to assess added value.

Kirkpatrick [24-26] gives a four level model of evaluation. The first level is how the students value the direct experience - also known as the smile sheet -, the second level measures the direct learning effect, the third level tests if a permanent learning effect has occurred and the fourth level looks for the added value in the workplace. The second level compares to the desire to evaluate the teaching effects. Kirkpatrick argues that measuring the second level can only be done if the first level is measured, and that the second level needs a pre- and posttest (criterion 1).

Using student assessment, they can be asked to assess their own ability to perform a task. Since students could be without experience, it would be better to ask students for their perceived ability (design criterion 2).

A weakness of SALG is the self evaluation factor as mentioned before. We will work on the premise that when self-assessment is used in a pre- and posttest, with the same scales, the only thing that is measured is the gain. This does require experience of the student in the subject of self assessment. If a student has never made a planning for a project before, the pre-self-assessment of a student's ability to make a planning could give inaccurate results. We have to compensate for this effect, which has to do with the students relative position on four stages in the learning cycle [27]. One usually starts in stage one: unconscious incompetent. The next stage is described as conscious incompetent, also referred to as 'you know that you don't know'. The following stages are conscious competent and unconscious competent. Going from stage one to stage two is learning, detecting this in a pre- and posttest on perceived ability poses a challenge, since it could cause the perceived ability to decline. The design of an instrument to test added value will have to account for this effect (design criterion 3).

Another challenge is the effect of students not completing the whole course and therefore not handing in a posttest. A level of attrition of 25% is not uncommon [7], which poses a challenge for the evaluation of the results. It is argued that given described uncertainties in self assessment it is unwise to compare pretest and posttest results without accounting for the attrition (design criterion 4).

Concluding, our design criteria for using SALG to assess added value are:

1. Use pre- and posttest
2. Use perceived ability
3. Account for learning stage one
4. Account for attrition

The effect of the learning cycle can be countered by asking the student for their perceived ability in the pretest and asking on what experience this perceived ability is based. Posttest results of students with no previous experience should be scrutinized for the reported difference in their perceived ability. To achieve this, it should be possible to link the pre- and posttest on an individual basis.

Linking the pre- and posttest on an individual basis provides an opportunity to account for attrition. Pretests of students without posttests should be discarded for the evaluation of added value.

The design proposal for using of SALG is therefore: a pretest which asks students about their previous experience with the subjects and asks their perceived ability on thes subjects. A posttest asks students again for their previous experience (in order to link them) and their (new) perceived ability on the subjects.

3. The case of a professional Master in Project and Process Management

All tests have been performed on the first seven months of a professional Masters program (420 student study hours), in which only students with relevant working experience are allowed. The first part of the Master of Project and Process Management in the Green Environment at Van Hall Larenstein focuses on soft skill development with coaches and actors, lectures on project management subjects and a complex group project. The complexity of this project is created by having the group find a project sponsor which was willing to let them work on a 'non undisputed problem'. In other words, the project has to involve some kind of controversy like conflicting stakeholders or failed first attempts. The soft skills development is on individual and group basis, particularly on personal effectiveness. The lectures on project management subjects host a mix of scientific publications, project management methods, process management approaches, theoretical exercises and group reflection on the group project relating to the theory.

Looking at the primary goals of the Master, a fitting performance test would be if students would be able to handle stakeholder conflict situations (of a professional nature), to produce project plans which would contain acceptable steps for all stakeholders, to execute and close a project according to plan while performing adequate risk management and to be able to build and maintain a project team. This description explains the choice of 'non disputed problem'. It also provides clues why a performance assessment as a pretest would be difficult in areas like executing a project (requiring time), adequate risk management, handling conflict situations (both requiring realistic situations) and building teams (requiring time and a team).

The first case study had ten students, working in two teams on separate projects with nine students finishing. The second case study had seven students with seven finishing, working in one team. The third case study had eight students split on two projects, number of finishing unknown yet.

The students met every week for skills training and lectures for two consecutive days, and were asked at the end of every day to reflect on the quality of the day (like: quality of the teaching staff, quality of interaction etc.), therefore accounting for the level 1 measurement. Teaching staff was informed of their performance and if necessary changed their approach based on these level 1 measurements. A short level 2 measurement was performed at the end of each lecture day by asking each of the students to describe in a few words what they learned. The teams also met for meet twice for a whole week to work on the project.

Since the primary author was a lecturer in this Master, there was a stringent focus on anonymity of the students participating in the research to avoid student and or researcher bias. Assessment forms were handed out and collected by others than the researcher.

4. Testing design one

The survey pretest asked for experience (working in projects, leading projects, sponsoring project), trust in knowledge (theory, methods and techniques) and trust in skills (starting a project, executing a project, bringing a project to an end and closing a project). The trust questions were given a five point Likert scale ranging from no trust to high trust. The experience was given a four point Likert scale ranging from none to experienced (more than five projects). When experience was claimed the student had to fill in the turnover of the most recent project. The pretest was handed out and filled in at the kickoff session of the program.

The posttest survey asked the same questions, and asked students where they perceived their learning gains (subjects project communication, leading projects, planning and monitoring projects, people skills, risk management, negotiation skills, decisiveness, team building, industry specific and project management methods and techniques based on the top ten competences derived from research [28]), in a three point Likert scale ranging from no progress to considerable progress. It also hosted a not applicable option. Industry specific means knowledge in their own domain, a 'control question' since specific attention on this subject was not featured in the courses. The posttest was supposed to be handed out and filled in at the closing session of the first part of the program, but was sent by e-mail afterwards instead.

The survey forms were not coded to an individual student, but asked their experience twice. Although the sample was small, it was not possible to link all posttest surveys to the pretest surveys. Six students returned the posttest, four could be linked to the pretest. We briefly describe their results.

Student 1: experienced project manager. This student reports a gain in trust on all three knowledge areas with one point and on two of the four skills, also with one point. This student reports some perceived progress on leading projects, own domain and project management methods and techniques) and considerable progress on people skills and team building. There was no perceived progress on the other subjects.

Student 2: experienced project manager. This student reports a gain in trust in knowledge, theory but a loss in trust in knowledge of techniques (both one point). No gain or loss was reported on trust in skills. The student did not complete the perceived progress part of the survey.

Student 3: inexperienced project manager. This student reports a knowledge trust gain on theory and on skills in executing a project (both one point). On perceived gains, this student reports no progress on negotiation skills, but some or considerable progress on all other subjects (considerable on project communication, planning and monitoring and own domain).

Student 4: experienced project manager. This student reports knowledge gains on all three knowledge areas (one point) and a loss of trust in skills on starting a project (from highly confident to confident). This student reports a perceived progress on all subjects, some progress on planning and monitoring and on decisiveness, the rest is reported as considerable progress.

No student reports a big change in their trust level. The reported changes could be attributed to different causes, with education being only one of them. It could also be respondent error, since asking for trust is a subjective measurement and small changes could occur depending on the respondents well being. The reported perceived gain shows a much more pronounced result than the trust questions.

Remarkable is that a experienced project manager (student 4) reports more perceived gains than the inexperienced project manager (student 3), which is in contrast to what could be expected. The causes of this unexpected result can only be speculated. Another experienced project manager (student 1) reports much less perceived gain.

5. Improved design and second test

Students numbers were introduced, linking pre- and posttest. The section of perceived gains was introduced into the pretest as a perceived ability. The perceived gains remained in the posttest, but a comparable perceived ability was introduced. The most important change however is the introduction of a 360 degrees feedback system. The students were evaluated by two peers and the coordinator and skills trainer (primary author is not the trainer). Students and trainers got acquainted with each other (that is a study week of working experience in their project) before this pre-test was taken. The self assessment pretest was taken in the first week of the program.

<insert figure 1>

There are seven participants in the Master. Two experienced students, two moderately experienced students and three students with little or no project management experience (although they all have project membership experience). The pre-test shows a that on average the students value their knowledge of project management theory, methods and techniques to be low while their peers and trainer value them much higher on the same scale. Lesser differences but sharing the same image are risk management, starting and executing projects. On the other questions there appears to be a congruent picture – on average. Almost all points of measurement are lower in the self-evaluation than that of peer or trainer evaluation (see figure 1 for an illustration).

Greater differences of level evaluation are mostly seen in the abstract first part of the questionnaire (theory, methods, techniques, starting, executing, ending & closing and the last question: knowledge of pmgt methods and techniques. Precisely in these parts the self-assessment reports the highest gains. The before and after comparison of the trainer and the peers show much more moderate differences with the initial appraisal. See figures 2a, b and c.

<insert figures 2 a, b and c>

It is tempting to conclude from these results that self-assessment is the least reliable instrument. Especially when noted that on average the project management methods and techniques get very little attention in the researched period. There is much more attention on interpersonal competences. Not graphically displayed, but comparing on the more abstract knowledge (theory, methods and techniques) and ability subjects (starting, executing, ending and closing), only self assessment shows a more than 'respondent error' effect. Self assessment is consistently more positive than peer or trainer assessment on these subjects.

Comparing claimed learning effect versus the difference in assessed levels, the average scores of the students show a high level of congruence, as do the scores of the peers. The scores of the trainer show less congruence. On three occasions the change in average confidence level of the trainer is almost zero (planning & monitoring, industry specific and project management methods & techniques) while the average claimed learning effect is high. The confidence assessment looks more reliable in this case as it has been argued before that last two fields receive very little attention so a high learning effect is improbable. This does not explain the difference on planning & monitoring though.

<insert figure 3>

Comparing the results of students, peers and trainers on learning effect or on difference (delta) on confidence levels, the latter shows agreement among the three groups. Which is strange because of the before noted temptation to discount self-assessment. The self-assessed learning effect and the peer-assessed learning effect also show agreement and a sort of disagreement with the learning effects claimed by the trainer. See figure 3 for an illustration.

The results are at least confusing. Looking at averages, the self-assessment appears to be the least reliable, but does show a reasonable agreement with peer- and trainer-assessment. The arguably less reliable learning effect claim shows better congruence albeit only between self and peer evaluation. This could support a conclusion that both are not reliable.

Looking from a different perspective, which student learned most, reveals that even on this level there is no agreement. Best student - the one with the highest 'added value' - according to self assessment is #4, while self claimed #5 is the best student, trainer assessments reveals #1 to be best which is tied best with #3 on trainer claimed. Peers assess #5 to be the best student and claim #3 to be the best student, as illustrated in figure 4.

<insert figure 4>

On average all measurements do suggest added value. This changes at student level, where several subjects show a negative assessed gain or a higher assessed gain than claimed and quite often disagreement between the assessed learning effect from different assessor groups (self, peer, trainer) ranging from negative to positive. These effects are not linked to specific subjects or specific assessor groups.

6. Tweaked design and third test

The third design in our research will add a few extra's in the posttest: It will feature a election of the best student: which of the students is (self, peer and trainer) assessed to be best in project management (ordering of all the students). This will be compared with an actual individual assessment of all students. This third design will again be tested on the Master of Project and Process Management, the pretest has been performed. This time the pretest has been issued not directly at the start of the course but after a few weeks. As usual, the peer and trainer assessment was done after the first project week. The average appraisal bears a great similarity with the average appraisal of the second test. The posttest will be performed in the beginning of July 2015. Although not expected, the results of this test could lead to changes in the discussion and conclusions session.

7. Discussion

The first test did not provide clues that any of the two types of measurements (SALG with difference in confidence level and SALG with claimed learning effects) could be a valid representation of learning gains. Introducing peer and trainer assessment did solve the issue of correctly linking pre- and posttest, but did not provide extra clues whether any of the measurements could serve as a valid representation of learning gains. Overall students and peers claim learning effects which are notably higher than comparing perceived ability.

The control question - industry specific - does lead to a negligent effect in the second test, only trainer claimed shows an average learning effect. If we take the small but positive effect as an indication of no learning, comparing the averages shows possible learning in self assessed on the topics of people skills, risk management and negotiating an even less likely on team building and leading. The only convincing topic being project methods and techniques. Because of this result, self assessment is discounted as method. Self claimed is discounted based on the results of the first test.

Claimed effects are almost consistently higher than assessed effects, not only by students self, but also by peers and notably by the trainer. The claimed effects are sometimes high on subjects where was expected nor logical (like project management methods and techniques and leading projects) and sometimes high on subjects where it was expected and logical (like team building and people skills). This leads to the conclusion that this kind of measurement does not measure what it is supposed to. This only leaves trainer assessed as a potential measurement, but especially this one shows inconsistencies with peer and self evaluation. And having a trainer assess the added value of his or her own actions does not seem right.

This is not a quantitative research. We only used a small sample to test whether the design criteria could remove the potential problems, like inaccuracy, attrition and stage one learning. And we used it on one course only. This may have been wise, as the averages appear to show some kind of learning, the student level reveals that this is mostly due to averaging out strange effects on student level. Tested on bigger numbers of students, this could have remained hidden.

The design used a relative rude scaling of the answer options, which allows students to quickly assess themselves and others. On the other hand it does not allow for small changes in confidence level to be detected. One can wonder if a student, peer or trainer would indeed give more reliable answers in a more detailed scale. Given the results, further research in this area seems pointless.

Our premise that comparing perceived ability would solve over- and underestimation cannot be proven.

8. Conclusions and suggestions

Our research set out to define criteria - based on literature - for using SALG as a method of measuring added value. The test revealed no clues that a design based on those criteria could be used as a method of measuring learning gains, not for use as a self assessment, peer assessment or trainer assessment of learning gains.

Our results supports earlier an earlier statement [23] that using self evaluation does not lead to results that can be trusted, with the addition that we also did not find any clues that a variant of 360 degrees assessment could provide trusted results.

This results underlines that 'Increase knowledge is relatively easy to measure /…./ we can measure attitudes with a paper-and-pencil test /… / [for skills] a performance test is necessary' [26]. Claiming teaching or learning effects using SALG in any form, even using 360 degrees feedback, does not appear to be valid. Not in the field of project management, but most probably in any field.

It would be interesting to disprove our findings. We do suggest that a better track for researching added value is the design of easy to administer but valid assessments of real ability. Further research on more easy to assess abilities could support our findings as we will try in the third test.

References

[1] Crawford, L.; Morris, P.; Thomas, J.; Winter, M. Practitioner Development: From Trained Technicians to Reflective Practitioners. Int. J. Project Manage. 2006, 24, 722-733.

[2] Pant, I.; Baroudi, B. Project Management Education: The Human Skills Imperative. Int. J. Project Manage. 2008, 26, 124-128.

[3] Martin, A. A Simulation Engine for Custom Project Management Education. Int. J. Project Manage. 2000, 18, 201-213.

[4] Ellis, R.; Thorpe, T.; Wood, G. E-Learning for Project Management. Proceedings of the Institution of Civil Engineers. Civil Engineering 2003, 156, 137-141.

[5] Michel, C.; Prevot, P. Knowledge Management Concepts for Training by Project; an Observation of the Case of Project Management Education. In International Conference on Knowledge Management and Information Sharing, Madeira Portugal.

[6] Nijhuis, S.A. Learning for Project Management in a Higher Education Curriculum. In Project Management Institute Research and Education Conference 2012, Limerick, Ireland.

[7] Lim, B.; Hosack, B.; Vogt, P. A Framework for Measuring Student Learning Gains and Engagement in an Introductory Computing Course: A Preliminary Report of Findings. Electronic Journal of e-Learning 2012, 10, 428-440.

[8] Beaubier, E.W.; Thayer, A.N. Project Leadership. Assessment Alternatives, 1973.

[9] Edwards, A.; Knight, P. Assessing Competence in Higher Education.; Kogan Page: London, 1995.

[10] Youker, R. Project Manager Success Criteria. PM World Journal 2012, 1.

[11] Morris, P.W.G.; Crawford, L.; Hodgson, D.; Shepherd, M.M.; Thomas, J. Exploring the Role of Formal Bodies of Knowledge in Defining a Profession – the Case of Project Management. Int. J. Project Manage. 2006, 24, 710-721.

[12] Turner, J.R.; Müller, R.; Dulewicz, V. Comparing the Leadership Styles of Functional and Project Managers. International Journal of Managing Projects in Business 2009, 2, 198-216.

[13] Ewell, P. An emerging scholarship : a brief history of assessment. In Building a Scholarship of Assessment.; Banta, T.W., Ed.; Jossey-Bass: San Francisco, CA, 2002, pp. 3-25.

[14] Kahneman, D. Thinking Fast and Slow.; New York Farrar Straus & Giroux 2011, 2011.

[15] Kuntze, A.J.(. Assessing Progress in Mastery of Counseling Communication Skills. 2009.

[16] Axelrod, R. Assessing Learning Gains in Political Science. Teaching Political Science 1976.

[17] Andriessen, D.; Aken, J. Handboek Ontwerpgericht Wetenschappelijk Onderzoek;Wetenschap Met Effect.; Boom Lemma: Netherlands, 2011.

[18] Seymour, E.; Wiese, D.; Hunter, A.; Daffinrud, S.M. Creating a Better Mousetrap: On-Line Student Assessment of their Learning Gains. In National Meeting of the American Chemical Society, San Francisco, CA., 27-03-2000.

[19] Vogt, G.; Atwong, C.; Fuller, J. Student Assessment of Learning Gains (SALGains). Business Communication Quarterly 2005, 68, 36-43.

[20] Anderson, D.; Burns, S. One-Minute Paper: Student Perception of Learning Gains. COLLEGE STUDENT JOURNAL 2013, 47, 219-227.

[21] Anderson, A.K. An Assessment of the Perception of Learning Gains of Freshmen Students in an Introductory Course in Nutrition and Food Science. Journal of Food Science Education 2006, 5, 25-30.

[22] Rooij, S.W.v. Scaffolding Project-Based Learning with the Project Management Body of Knowledge (PMBOK®). Comput. Educ. 2009, 52, 210-219.

[23] Symons, A.B.; Swanson, A.; McGuigan, D.; Orrange, S.; Akl, E.A. A Tool for Self-Assessment of Communication Skills and Professionalism in Residents. BMC Medical Education 2009, 9.

[24] Kirkpatrick, D. Four-Level Training Evaluation Model. US Training and Development Journal 1959.

[25] Kirkpatrick, D. Great Ideas Revisited. Techniques for Evaluating Training Programs. Revisiting Kirkpatrick's Four-Level Model. Training and Development 1996, 50, 54-59.

[26] Kirkpatrick, D.L.; Kirkpatrick, J.D. Evaluating Training Programs the Four Levels. 2006.

[27] Maslow, A.H. Motivation and Personality.; Harper & Row: New York, 1954.

[28] Nijhuis, S.A. Project Management Competences to Incorporate in a Higher Education Curriculum. FOR 2012, 90, 115-120.

Authors

Principal Author: Steven Nijhuis holds a masters degree in Mathematics from the Technical University of Eindhoven. After practicing project management and consultancy in the steel making industry and built environment he is currently a researcher of learning project management and a project manager in higher education.

Co-author: Ruben Vrijhoef holds a PhD from the Delft University of Technology. He is a Professor of Applied Sciences at the Utrecht University of Applied Sciences in the field of supply chain redesign in

the built environment. He is also a senior researcher at the Delft University of Technology in the field of construction management, supply chain management, lean construction and construction logistics.

Co-author 2: Dr. Joseph Kessels is professor of Human Resource Development at the University of Twente. Until 2008 he served as Dean of TSM Business School. Joseph Kessels has a specific research interest in the characteristics of learning environments that support knowledge productivity and that facilitate innovation. Distributed leadership is a more recent research topic, specifically related to professional development.

Tables and figures

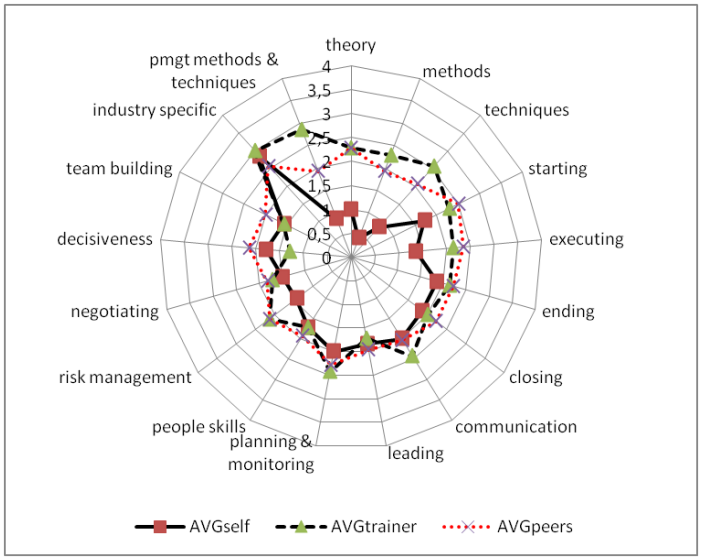

Figure 1. Comparing initial self, peer and trainer appraisal

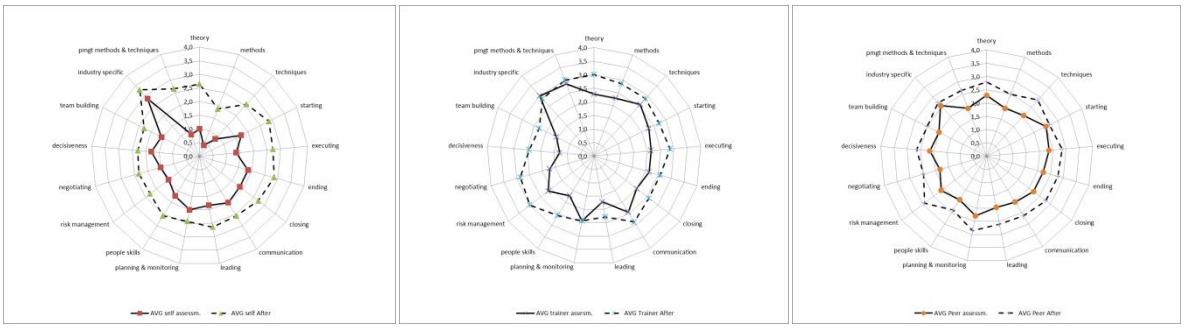

Figure 2a, b and c: comparing pre and post from self (a), trainer (b) and peer assessment (c)

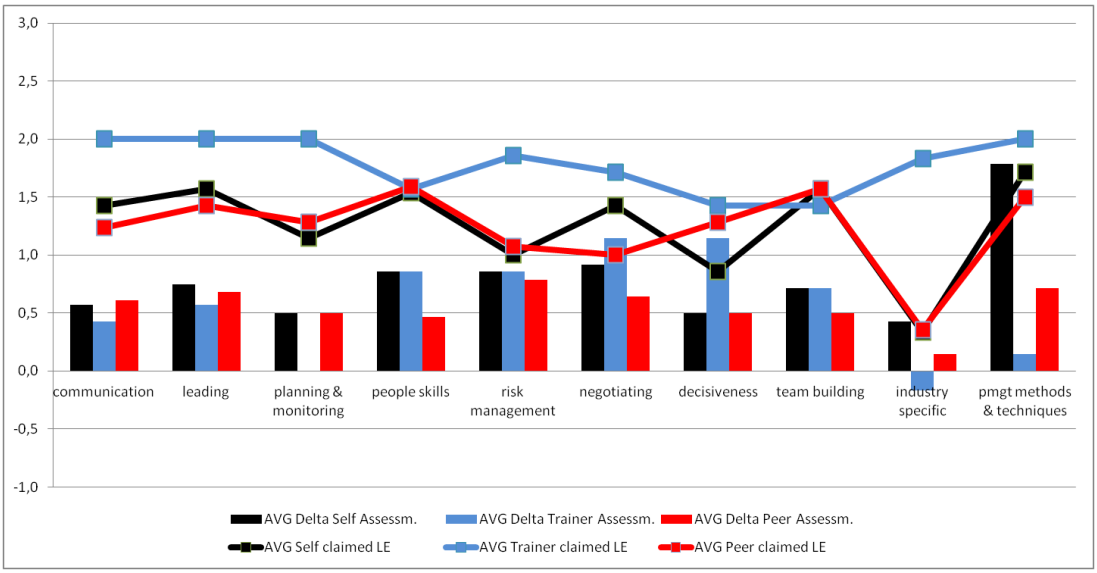

Figure 3. Comparing averages assessed and claimed learning effects on subjects

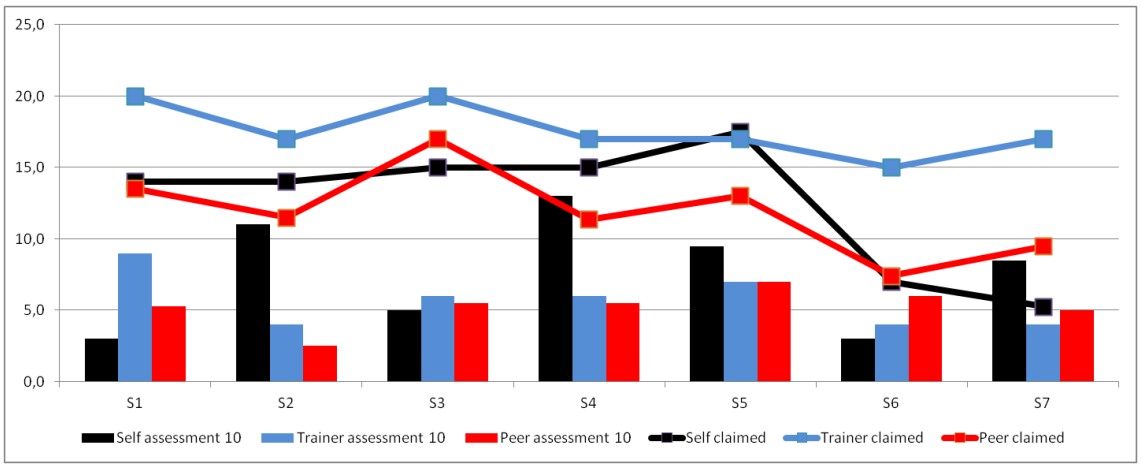

Figure 4. Comparing averages assessed and claimed learning effects on student level