ASSESSING ESSAYS TO DEVELOP WRITING

Peter Karlsudd
Linnaeus University, Kalmar, Sweden
E-mail: peter.karlsudd@lnu.se

Abstract

Special needs teacher education in Sweden has by tradition been pragmatic and action-oriented. The higher degree of academization the education is now undergoing has revealed deficiencies in the concluding essay project. A low problematization degree, poor grounding in theory as well as undeveloped analyses are among the explanations of the low quality.

This article describes a development project conducted at Linnaeus University with a view to improving the quality of student essay writing. With the overarching aim of making students aware of the quality criteria involved in essay writing, students and teachers assessed essays on the basis of 24 different criteria. Initially worked out by teachers, the criteria were developed and supplemented in dialogue with students. After the appraisal the conformity of assessment between student-student, teacher-student and teacher-teacher was presented.

The result demonstrates that disagreement exists in the assessment. Teachers and examiners assess the quality of the essays as being somewhat lower in certain respects than what students do, whereas they may give higher values on the basis of other criteria. However, individual differences in assessment among teachers are not fewer than among students. There are clear differences in the evaluation of method, theory grounding, innovativeness and creativity. Subsequent discussions have led to clearer quality requirements for essays and have also to some extent reduced the former interpretation gap.

Key words: *assessment, assessment system, essay writing, essay seminar, teacher education.*

Background

The education of teachers and special needs teachers in Sweden has in the last decade become more and more academically oriented. The reason is that teachers-to-be, including special needs teachers, have not been given a real chance to develop their abilities to adopt a questioning and reflective attitude. It is essential for a scientific and critical approach to clarify the foundation of different lines of argument and bring hidden assumptions to light. This also includes the ability to distinguish between personal views and better founded-arguments. This attitude is of great professional use to all teachers. Their work must rest on scientifically based arguments and not solely on common sense, tradition or established habits. It is not unusual for such concepts to characterize non-scientific and unreflective acts.

The Swedish Government Official Report "En hållbar lärarutbildning" (Sustainable teacher education, SOU 2008:109) lays stress on a scholarly and critical approach and proposes reinforcing theory of science, research methodology and statistics in a new teacher education. There is reason to believe that similar measures will be proposed for the further education offered in the special needs teacher programme.

Previous studies have demonstrated that scholarly articles are seldom used by Swedish teacher students. In an evaluation of degree projects carried out in teacher training programmes

Forsberg and Lundgren (2006) stated that the mean value of references in degree projects is 22 references with an average of 3 scholarly articles/research reports per project. The result of the study indicates a positive connection between the quality of the essay and the number of references, including references to works in other languages than Swedish. A study by Lyngefelt (2009) establishes that course literature of a textbook character has been used to a large extent. The authors maintain that the students' motivation to use scholarly literature decreases when the course literature consists of textbooks. Against this background the above-mentioned report advocates a consumption model for degree projects, which entails that students conduct literature studies, i.e., research surveys of published studies. This procedure will make students into research consumers rather than research producers. "In this way future teachers are prepared for a continuous use of current research" (SOU 2008:109, p. 201). Research-based knowledge will enable teachers to work more professionally to the benefit of students and schools while simultaneously endeavouring to regain the status once held by the teaching profession (Säljö & Södling 2006).

If at an early stage of the education guidelines can be provided on what sources of information are trustworthy, the students' chances of doing a good degree project increase (Davis, 2003). For this reason it is advisable that teacher students develop their knowledge of what types of publication are attributed a high value within their own subject area. Increasing one's knowledge of information sources of different kinds contributes towards a higher estimate of doctoral theses and scientific publications (Chu & Law, 2007). Using quotation analysis and reference analysis may also promote information search and reference work as well as facilitating the understanding of the structure of a scholarly work (Brazzeal & Fowler, 2005). One way of increasing this understanding is to provide students and teachers with the opportunity to examine and discuss such work from the starting point of common well established quality factors.

Both the interaction between teachers and students (Jokela & Karlsudd, 2007; Bennett, 1998) and teachers' understanding of students form a highly important quality factor for learning. It rests with the teacher to capture the students' previous knowledge and make room for their own choices and tasks (Christensen & James, 2001). What is right and self-evident for one individual may be totally unintelligible to another one, and understanding what others mean involves trying to see the world from their point of view (Booth & Marton, 1997). To create coherence and continuity in teaching and learning also requires cooperation among the teachers and supervisors involved in the students' learning (Tucker & Bryan, 1999). Students use different learning strategies (Entwistle & Ramsden, 1983; Marton, 1999), and offering alternative ways of expression will benefit the learning environment. It is important to make students face problems and tasks that feel meaningful and are neither too easy nor too difficult to solve (Yorke, 2004). For this very reason it is essential for teachers and supervisors to obtain a clear picture of the students they meet professionally. One way of turning students into more active research consumers is to make them partake of previous student degree projects and discuss quality on the basis of shared and familiar concepts. This article describes a development project where a number of quality aspects have been the focus of assessment and discussions.

Purpose

The purpose of the present project is to make use of a systematic close reading and assessment of various doctoral theses and an approved degree project produced jointly to make students aware of the quality criteria demanded of a degree project and thus improve the quality of essay writing. Further, teachers and supervisors are expected to create a clearer picture of how students react to various assessment criteria to enable them to better adapt their teaching and supervision.

Methodology of Research

The method applied in the project is action research. It could be described as an interactive transformation process between action, research, theory and practice, with the researcher taking an active part in the process (Starrin, 1993). The action research approach has not been uniformly formulated, since different researchers focus on different aspects. Different terms circulate in an attempt to arrive at a better definition of the approach (Tiller, 1995). Argyris (1985) uses the term action science for the part of action research that is more theoretically oriented, and action research for the practically oriented part. The most common type of action research is when, as in this study, students, supervisors and examiners work together with the researcher (Cohen & Manion, 1994). It is not unusual for critical thinking and new issues to emerge from this work procedure (Feldman & Minstrell, 2000; Archer, Holly & Kasten, 2001). This article describes how students and teachers together assess a degree project in order to increase the understanding of the qualities, which form the basis of grading. Rust, Price and O´Donovan (2003) discuss the possibilities of improving the examination process by increasing the understanding of assessment criteria. Despite the rather limited elements used to clarify these criteria, the results indicate an improvement in student performance.

Starting Points for Assessment

In this study students were asked to read and assess a doctoral thesis of their own choice. After a discussion of the quality of the theses students and supervisors read an essay that had obtained the "Pass" grade in an earlier course. The assessments were conducted with support of an assessment template divided into three levels, termed frame, external structure and internal structure (Figure 1). The template was designed by teachers with a PhD and was developed in parts and supplemented in dialogue with the students. With the help of 24 criteria presented in Table 1 and explained below students and teachers assessed the quality of the essay.
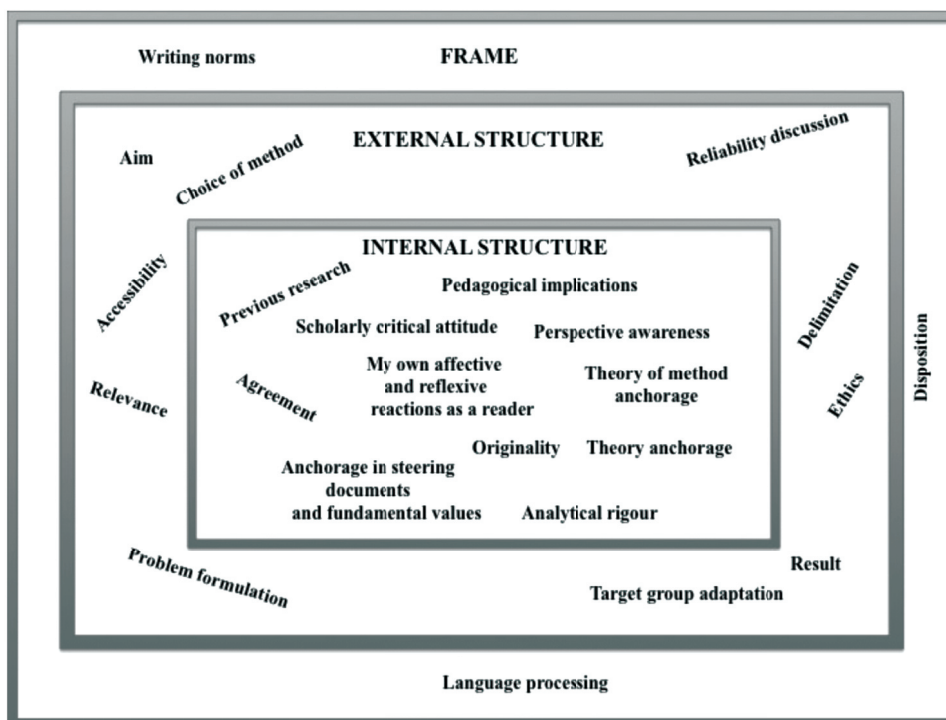
Figure 1:   The three essay quality levels: frame, external structure and internal structure.

Frame:

The first quality forming the basis of the assessment is defined as "frame" and consists of three assessment criteria. The first criterion is writing norms (1), which include the ability to handle quotations, references, figures and tables correctly. The second is language processing (2), which involves the requirement that the language should be varied, distinct and correct so as to guide the reader through the text. The third is disposition (3), by means of which the reader will encounter a balanced essay containing a clear connection between research question, previous research, theory, result, analysis and discussion. The author should also avoid a too extensive presentation without losing variation and depth.

External structure:

The second essay quality is termed external structure and consists of ten assessment criteria. The first criterion is aim (4). What is assessed here is whether the author has managed to choose research questions that are possible to examine. Is the knowledge that is sought after well defined and the choice of subject motivated? The following one is delimitation (5). This determines whether the research question is clearly delimited and whether the author's perspective is made clear in the text. The next item, problem formulation (6), also checks whether the research question is clear and stringent, preferably expressed in the form of hypotheses or questions. The accessibility item (7) assesses the quality of the structure and whether the concepts are clearly defined. A closely related assessment criterion is target group adaptation (8). This is where the text is judged for its readability and intelligibility to the target group. In result (9) the ability to process, analyze and interpret the research material clearly is assessed. The result should appear from a lucid and logical presentation. In relevance, the ensuing item (10), the contents of the essay are tested against their importance to the education context. Whether anything important has been left out or anything unnecessary has been included are questions that may be posed. The ethics item (11) assesses whether the author abides by the ethical rules set up and whether the work itself bears witness to a well-founded ethical awareness. Choice of method (12) is the item where the relevance of and motivation behind the choice are scrutinized. The last criterion concerning external structure quality is the reliability discussion (13). This is where the author's arguments for the credibility of the result are evaluated. Discussions about validity, reliability, missing values and generalizability are among the concepts focused on in this criterion.

Internal structure:

In the third quality the first assessment criterion is perspective awareness (14). Here the author is expected to declare his pre-understanding and make an account of the research situation. This is the item where the author's own position is stated. Here am I! This is where I start from! Under the previous research item (15) the author's ability to present current research is assessed. It must be clearly connected with previous research and the author has to compare different schools with one another. The interpretation theory applied is evaluated under theory anchorage (16). Have the chosen theory/theories contributed to a deeper understanding of the subject under study, and does the author argue well for the choice of theory? The next item, theory of method anchorage (17), concerns whether the author's choice of method is motivated by a detailed account clearly linked to the literature on methodology. Analytical rigour (18) is a measure of whether the author's text is analytical, pithy and rich in contents. This item also includes assessing whether theoretical concepts are adequately applied in the analysis. When the scholarly critical attitude (19) is to be evaluated the author's critical approach is scruti-

nized. Are the conclusions drawn from the result well founded and coupled to previous research and theory? Does the author discuss any alternative conclusions? The agreement (20) between problems, method and results is assessed in the following criterion. In the presentation of results the author is to provide answers to the problems formulated as well as present conclusions, followed by a chapter containing the discussion of the results. For evaluating the novelty value and originality of the essay there is an originality criterion (21). This investigates whether the author has given evidence of novel thinking in formulating aim and questions. Creative thinking in, for example, the choice of method is also considered. The consequences the essay may entail for future practice as well as its possible contribution to research and development are estimated in the item termed pedagogical implications (22). The last but one criterion is denoted anchorage in steering documents and fundamental values (23), which measures links in the essay to current steering documents and to the fundamental values stated for present and future activities The last item is expressed as my own affective and reflexive reactions as a reader (24). For the text to obtain a high grade it is supposed to trigger thought processes, to involve, affect and arouse the reader's interest and to make an impact. The third quality level ends with an item for any other comments.

The Study Implementation

To start with, the students read a thesis of their own choice, which was then discussed on the basis of the specially designed assessment template. Thereupon discussions were held about the 24 criteria. On a later occasion the students read a joint essay that had obtained the Pass grade in a previous course. After that the students assessed the essay individually, following the same template. When the individual assessments were completed the result was compared and discussed in small groups. The group discussions were documented and the written material together with the assessment and the comments made in the questionnaire formed the basis of the presentation of results. An important part of the exercise was that the students had to give constructive suggestions for amendments to the work they had read.

The teachers/supervisors involved in the course independently of each other assessed the same essays as the students without knowing the result of the assessments made by the others. This was followed by a short interview of about half an hour. In order to check agreements and disagreements the mean value and standard deviation of the questionnaire result were applied.

The assessment was made according to an interval scale from 1 to 5, where 1 represented the lowest and 5 the highest grade. In choosing a descriptive measure, such as a central or distribution measure, the distribution curve of the observation material is of great importance. If the approximate distribution of the measurement values is normal, i.e., if most observations end up in the middle and there are few extreme-value observations, arithmetic mean and standard deviation are used (Löfgren, 1991). In the data collected with support of the student group's assessment form these measures were deemed to be appropriate, since the distribution was normal. As the number of teachers was small, this principle cannot apply to that group, but presenting the result with support of the same statistical measures will, however, form a basis for comparison. All the results are only valid for this project and its population and cannot be statistically representative of any other sample. When describing standard deviation a value of 0 makes no difference in the assessment while a difference closer to 1 indicates a definite variation. In the assessment the students differed at the most by 4 steps on the scale, whereas the teachers' assessment never differed by more than one step.

Results of Research

The results of the student assessments are relatively uniform. The assessment criteria where the students diverge the most are originality, pedagogical implications and affective and reflexive reactions (Items 21, 22 and 24, Table 1), The students differ in their opinion about what is original, of benefit to present and future activities, and about what made them react emotionally and intellectually to the text. One explanation may be the difference in students' specialization and interest. Some students' training focuses on younger pupils, some on older ones, with notable variations in interest and subject area as a consequence. The originality, pedagogical implications and affective and reflexive reactions items are probably more difficult to formalize and were felt by many students to be the most subjective assessment values.

Quite a few student assessment values lie close to one another. One example is previous research and theory anchorage (Items 16 and 17). A great many students place the summary of previous research on a par with links to the theories and models expected to provide a blueprint for explaining the result. On that score the teachers' assessments differ in that they were capable of elucidating the distinction between the two criteria.

Table 1.   Students and teachers' assessments of the essay.

	Arithmetic mean Students N.36	Arithmetic mean Teachers N.3	Standard deviation students N.36	Standard deviation teachers N.3
FRAME				
1) Writing norms	3.36	3.00	0.64	0.00
2) Language processing	3.06	2.67	0.58	0.58
3) Disposition	3.36	3.33	0.68	0.58
EXTERNAL STRUCTURE				
4) Aim	3.14	3.00	0.49	0.00
5) Delimitation	3.11	2.67	0.46	0.58
6) Problem formulation	3.28	2.67	0.72	0.58
7) Accessibility	3.22	3.00	0.68	0.00
8) Target group adaptation	3.28	2.67	0.68	0.58
9) Result	3.11	3.00	0.60	0.00
10) Relevance	3.50	3.33	0.56	0.58
11) Ethics	3.19	3.00	0.40	0.00
12) Choice of method	3.03	3.33	0.61	0.58
13) Reliability discussion	3.11	3.00	0.67	0.00
INTERNAL STRUCTURE				
14) Perspective awareness	3.39	2.67	0.55	0.58
15) Previous research	3.42	3.33	0.69	0.58
16) Theory anchorage	3.14	2.00	0.68	0.00
17) Theory of method anchorage	3.11	2.67	0.71	0.58
18) Analytical rigour	3.11	2.67	0.71	0.58
19) Scholarly critical attitude	3.31	2.33	0.75	0.58
20) Agreement	3.03	3.00	0.74	0.00

21) Originality	2.83	3.33	0.94	0.58
22) Pedagogical implications	3.08	3.33	0.91	0.58
23) Anchorage in steering documents and fund. values	3.22	3.00	0.59	0.00
24) My own affective and reflexive reactions as a reader	3.06	3.33	1.19	0.58

When the teachers discussed the three qualities, the third quality was given the highest value. This is not self-evident in the students' assessments. Many of them focus on the frame, in other words on the more immediate and visible qualities of the report. By a quick glance it is fairly easy to assess the first three items, and students who are inexperienced in essay writing concentrate a great deal on external qualities. One may argue about the relative assessment of these three levels, but all the teachers agreed that a careful reading of the third level provides a better basis for judging the level of the essay and for grading.

A tangible experience from discussions with the students was that a number of penetrating and constructive suggestions for improvements were presented. None of the students awarded the "joint essay" a Pass with Distinction grade, and several of them gave it a Fail grade. There was extra space at the bottom of the questionnaire for further comments. It emerged that certain questions were thought of as overlapping one another and that some of them felt particularly hard to assess. One reflexion from the teachers involved in the course was that the students' analytical and critical ability was often greater when examining the essays of others than when analyzing their own work.

Pedagogical Consequences

During the essay seminar:

The result shows that many of the students displayed penetrating and constructive suggestions for improving the essay they had read. One immediate consequence of this was to give students a greater responsibility in the assessment and grading process that followed their own and their fellow students' degree projects. In order to give greater weight in grading to criticism and defence, the assessment of the opponents' criticism and the respondents' defence was formalized. The opponent was then assessed on the basis of three criteria, the first of which implies whether the opponent was well prepared and dealt with the work according to the recommendations given beforehand. The second criterion is that the opponent problematized and discussed, and the third that the criticism was constructive and contained suggestions for improvement. The respondent's acting is also assessed, whereby the first criterion is explaining and clarifying, the second responding to the criticism by a research-based attitude, and the third being open to criticism and using the opponent's suggestions to improve the work.

After the seminar:

When the discussion of the essay was concluded a private talk was held between essay writer, examiner and supervisor on how they experienced it all. After that the writer drew up a list of, primarily, the opponent's views but also included suggestions for changes and improvements from the examiner and the participants in the seminar. In the list of changes the writer is expected to itemize the suggestions and then argue item by item whether he or she is willing

to make changes or chooses to refrain from such. The respondent can thus take a stand against making the changes suggested by the opponent. Then the minutes are submitted to the examiner, who approves and sometimes deletes changes or suggests new ones. In this way the students largely formulate their own changes and participate more actively in the decisions taken with regard to completing the essay. The project result shows that the students' suggested changes are often more comprehensive than the demands made by the examiner before the system with the correction list started.

Discusssion

By getting closer to the students' horizon of understanding and acquiring a better insight into how they judge the quality criteria which form the basis of assessing scholarly work the quality of the students' own essays will probably increase. This agrees well with Christensen and James' study (2001). Reducing the number of traditional course handbooks and introducing scholarly articles and reports early on in the programme should be an efficient way towards a higher goal fulfillment (Davis, 2003). By clarifying goals and assessment criteria and making students to some extent cooperate in specifying and formulating the criteria the quality should improve. This is the experience of previous studies (Bloxham and West, 2004). One suggestion for a future project would be for students and supervisors to work together in research projects where the problem has been formulated by students. This would probably increase work equality in comparison with the present situation, albeit to a small extent, where supervisors formulate the problem directing the work towards their own research.

As for the credibility of the result it may be assumed that the students felt relatively uncertain in their assessment. If the assessment method described in this article becomes properly anchored in essay courses, the dispersion noticeable in some respects might decrease.

Conclusions

In sum, it may be stated that the quality of degree projects will improve if students and teachers discuss the assessment criteria together. To involve students in the assessment and take better note of the improvements suggested by the opponents has turned out to be a successful method. To jointly discuss quality on the basis of a well-planned assessment system will also improve student understanding of the quality criteria behind scholarly work. It will also give supervisors a better insight into students' previous knowledge and pre-understanding of essay writing. Using the same assessment template in the course provides participants and course organizers with a language in common which facilitates interpretation and assessment. One suggestion that has not been tested yet is attributing a higher value to the third internal-quality level by, for instance, multiplying this value by two. This might reduce the emphasis on the form of the essay in favour of spending more effort on elements like analysis and theory anchorage. One concrete result is that the distinction between previous research and theory of science has become clearer after the joint exercises. The seminar discussions have improved and become more constructive and the students are better motivated in their revision work, since the changes are largely their own suggestions.

References

Archer, J. M., Holly, M.L. & Kasten, W.C. (2001). *Action Research for Teachers*. Upper Saddle River. NJ: Merill/Prentice Hall.

Argyris, C., Putnam, R., McLain Smith, D. (1985). *Action Science*. San Francisco: Jossey Bass.

Bennett, J.B. (1998). *Collegial Professionalism. The Academy, Individualism and the Common Good.* American Council on Education: Series on Higher Education: Phoenix, Arizona: Oryx Press.

Bloxham, S. & West, A. (2004). Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment & Evaluation in Higher Education,* Vol. 29, No. 6, pp. 721–733.

Booth, S. & Marton, F. (1997). *Learning and Awareness.* New Jersey: Erlbaum.

Brazzeal, Bradley & Fowler (2005). Patterns of information use in graduate research in forestry: A citation analysis of Master's theses at Mississippi State University. *Science and Technology Libraries,* Vol. 26, pp. 91–106.

Christensen, P. & James, A. (2001). "What are schools for?" The contemporal experience of children's learning in Northern England. In L. Alanen & B Mayall (Eds.): *Conceptualizing Child-Adult Relations.* London and New York: Routledge Falmer, pp. 70–85.

Chu, Samuel Kai-Wah & Law, Nancy (2007). Development of information search expertise: Research students' knowledge of source types. *Journal of Librarianship and Information Science,* Vol. 39, pp. 27–40.

Cohen, L. & Manion, L. (1994). *Research Methods in Education.* 4th edn. London: Routledge.

Davis, Philip M. (2003). *Effect of the web on undergraduate citation behavior.* Portal: Libraries and the Academy, Vol. 3, pp. 41–51.

Entwistle, N. J. & Ramsden, P. (1983). *Understanding Student Learning.* London: Croom Helm.

Feldman, A. & Minstrell, J. (2000) Action research as a research methodology for the study of the teaching and learning of science. In A.E. Kelly & R.A. Lesh (Eds.): *Research Design in Mathematics and Science Education.* London: Lawrence Erlbaum Associates.

Forsberg, E. & Lundgren, U. P. (2006). *Examensarbetet inom den nya lärarutbildningen: Tematiska studier.* Stockholm: Högskoleverket.

Jokela, P. & Karlsudd, P. (2007). Thesis web dialogue. *European Journal of Open Distance and E-Learning.* Issue 2007/I.

Lyngfelt, A. (2009). *Konsten att få "praktiska insikter i vetenskapligt tänkande": En studie om lärarutbildningens examensarbeten.* [The art of making "practical understanding of scientific thinking": A study of teacher education degree jobs]. In Hansson, B. & Lyngfelt, A. (Eds.): Pedagogiskt arbete i teori och praktik: Om bibliotekens roll för studenters och doktoranders lärande. [Educational work in theory and practice: If the role of libraries for students and doctoral students' learning]. Lund: Btj.

Löfgren, H. (1991) *Statistisk dataanalys.* [Statistical data analysis]. Studentlitteratur, Lund.

Marton, F. (1999). *Inlärning och omvärldsuppfattning.* [Learning and world view]. Stockholm: Prisma.

Rust, C., Price, M. & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education,* Vol. 28, No. 2, pp. 147-164.

Säljö, R. & Södling, M. (2006). *Utbildning på vetenskaplig grund: Röster från fältet.* [Education based on science: Voices from the field]. Stockholm: Högskoleverket.

SOU 2008:109. *En hållbar lärarutbildning. Betänkande av Utredningen om en ny lärarutbildning.* [Sustainable Education. Report of the Inquiry on a new teacher]. Stockholm: Fritzes.

Starrin, B. (1993). Tillämpad socialforskning [Applied Social Research]. In Holmer, J. & Starrin, B. (Eds.): *Deltagarorienterad forskning* [Participatory research] (pp.11-26). Lund: Studentlitteratur.

Tiller, T. (1995). Action learning and action research – opportunities and dilemmas. In Svein & Dowling (Eds.): *Reflections on Educational Research: The Qualitative Challenge.* Landås: Caspar Forlag A/S.

76

Tucker, A. & Bryan, R.A. (1999). *Charing the Academic Department Leadership among Peers*. Phoenix, Arizona: Oryx Press.

Yorke, M. (2004). Retention, persistence and success in on-campus higher education, and their enhancement in open and distance learning. *Open Learning,* Vol.19, No. 1, pp. 19–32.

Adviced by Vincentas Lamanauskas, University of Šiauliai, Lithuania

| *Peter Karlsudd* | Associate Professor, Linnaeus University, SE-391 82 Kalmar, Sweden. Phone: +46 772-28 80 00. E-mail: peter.karlsudd@lnu.se Website: http://lnu.se/employee/peter.karlsudd?l=en |