

To be published in *Assessment & Evaluation in Higher Education*

DOI: 10.1080/02602938.2012.751962

## **Evaluating the level of degree programmes in higher education: the case of nursing**

Trudy Rexwinkel\*, Jacques Haenen and Albert Pilot

Faculty of Social and Behavioural Sciences, Department of Education, Utrecht University, Utrecht, The Netherlands

### **Abstract**

The European Quality Assurance system demands that the degree programme level is represented in terms of quantitative outcomes to be valid and reliable. To meet this need the Educational Level Evaluator (ELE) was devised. This conceptually designed procedure with instrumentation aiming to evaluate the level of a degree validly and reliably still needed empirical evidence. In this study the ELE was carried out in collaboration with five professional bachelor nursing degree programmes. The degree programme level was conceptualised into disciplinary thinking and professional attitude. Disciplinary thinking refers to higher order thinking on complex disciplinary problems. Professional attitude refers to processing attitudinal characteristics of the profession that the study addresses. The level was operationalised into themes, resulting in good face validity: nursing science, care plan, diagnosing, and nurse-person relationships, and specified by learning outcomes representing the bachelor level (0.78). The learning outcomes were processed in a questionnaire for recent graduates (N = 470). Four components  $\geq 1$  eigenvalue measured the conceptualized themes explaining 72–80% variance and scale reliability from 0.78 in the single samples. With this study the level was evaluated validly and reliably and the conceptual design was demonstrated to be evidence-based.

Keywords: degree programme level; evaluation; construct validity; reliability

## Introduction

This study focuses on the problem of evaluating validly and reliably the level of degree programmes in higher education. This problem became crucial with the Bologna Agreement in 1999 when European Ministers of Education agreed to implement the bachelor and master structure and a system of quality assurance, the accreditation system. In accordance with the accreditation system, it is necessary to prove the real level with quantitative outcomes. These criteria were applied by the European Association for Quality Assurance in Higher Education (ENQA), which was mandated by the European Ministers of Education to implement a European Quality Assurance system in order to realise the system of accreditation. To facilitate this implementation, ENQA developed the European Standards and Guidelines (ESG) (ENQA 2009), in collaboration with the member countries. The national accreditation organisations elaborated these criteria in their frameworks. This has also been carried out in the Netherlands (NVAO 2011).

Policymakers from various European countries have developed the Dublin Descriptors to describe the levels of bachelor and master degrees (Joint Quality Initiative 2004). It was intended that these qualifications should be used in higher education in the European countries to enable cross-national comparison of the degree programme level. The policymakers agreed on the generally formulated descriptors. However, these descriptors are not based on peer-reviewed literature and neither are they empirically underpinned. It appeared to be difficult to conceptualise the degree programme level with the Dublin Descriptors and to underpin these empirically with valid and reliable descriptors needed to use them for cross-national comparison. The descriptors are worded in broad terms, allowing many interpretations, which lead to ambiguity and hinder the cross-national comparison of the degree programme level.

The Organisation for Economic Cooperation and Development (OECD) has explored the extent to which HEIs effectively prepare their students for participation in increasingly knowledge-based economies. The OECD Assessment of Higher Education Learning Outcomes project concluded that ‘in most countries, assessment results are inaccessible and [...] if HEIs would specify the expected student outcomes explicitly and in a measurable way, comparative assessment of learning outcomes would become feasible’ (Nusche 2008, 5).

Other initiatives have been undertaken and instruments have been developed, but these have found no broad commitment. For example, the Quality Assurance Agency (QAA) in the UK publishes subject benchmark statements, indicating the degree programme level,

which higher education institutions are expected to consult when designing, delivering and reviewing programmes. However, within the context of higher education, this is considered to be unacceptably bureaucratic and the empirical underpinning has not been demonstrated (Bellingham 2008). In the US also, systematic evaluation of the effectiveness of subject benchmark statements in assuring academic quality and standards is lacking and the goal of providing adequate evidence of student learning remains elusive (Ewell 2010, 173, 274).

In the discipline of Medicine, instruments are available that refer to the content needed to describe the level. These instruments focus e.g. on the level of preparedness for practice (Morrow, Johnson, Burford, Rothwell, Spencer, Peile, Davies et al. 2012). However, these instruments have been developed with the aim of evaluating specific aspects of the medical domain, instead of assessing the quality of the learning outcomes validly and reliably for the full degree programme level in a way that is applicable for a broad range of domains.

Problems that are related to evaluating are experienced by accreditation organizations. Various sources report that expertise in evaluation is too little or lacking in accreditation committees (Sandahl et al. 2006; Konrad et al. 2007; Crochet et al. 2009).

To summarise, evaluations that are based on valid and reliable outcomes lead to correct conclusions. To respond to this need, a procedure with instrumentation aiming to evaluate the degree programme level validly and reliably, the *Educational Level Evaluator* (ELE), was designed. The design of the ELE needed empirical evidence. Consequently the main research question is: *Can the degree programme level in higher education be evaluated validly and reliably by the Educational Level Evaluator?* This research question is detailed with three sub-questions: 1) How is the degree programme level conceptualised? 2) How is the degree programme level to be operationalised? 3) How is the degree programme level measured?

First, the concept of the degree programme level is elaborated. Then the method will describe the procedure to operationalise and measure this level. The results section reports the analysis and findings as reached by the *Educational Level Evaluator*. The paper closes with discussion.

### ***Conceptualising the degree programme level***

One of the most important aims of degree programmes in higher education is that students learn to think on complex disciplinary problems at the bachelor or master level. Higher education is the main place for students to develop the thinking capability that is considered

as the key of the degree programme level. This thinking involves a cognitive and an attitudinal component: disciplinary thinking and professional attitude.

Disciplinary thinking refers to higher order thinking on complex disciplinary problems. The level of academic challenge is one of the five benchmarks of the National Survey of Student Engagement (NSSE) in the US ([Kuh et al. 2001](#)). It refers to challenging intellectual and creative work that is central to student learning. Providing academic challenge concentrates on the nature and amount of the assigned academic work, the complexity of cognitive tasks, the use of higher order skills, and the standards that faculty members use to evaluate student performance. However, to understand better the processes and outcomes of providing academic challenge more information and specific data are needed ([Rutowski and Rutowski 2010, 411](#)). [Cosgrove \(2011\)](#) determines that critical thinking as well as a deep approach to learning lead to successful students. In a quantitative study of Oxford undergraduates (N=2330), he concluded (p. 344) that Oxford students with a deep approach to learning, which he defined as learning associated with critical thinking, are more successful in exams and felt more confident, supported and motivated than students who relied on superficial approaches to learning associated with uncritical thinking.

The degree programme level is conceptualised with the professional attitude too. This attitude refers to higher order processing of attitudinal characteristics of the profession or field the study is addressing, e.g., the integrity of the accountant or the independence of the scientist. This could be exemplified with the following description of a bachelor psychologist: “A bachelor psychologist is able to handle ethical dilemmas in the professional researcher’s field. He is able to handle these rationally and emotionally in correspondence with the professional code of psychologists” ([American Psychological Association 2012](#)).

The degree programme level is conceptualised as having a cognitive and an attitudinal component. The cognitive component refers to disciplinary thinking, implying higher order thinking on complex disciplinary problems. The attitudinal component refers to higher order processing of attitudinal aspects of the profession or field the study is addressing.

## **Method**

The *Educational Level Evaluator* (ELE) is a procedure with instrumentation aiming to evaluate the degree programme level validly and reliably. The *ELE* has been designed (Rexwinkel, Haenen & Pilot 2012) and currently needs empirical evidence. This study was carried out in collaboration with five degree programmes. In 2006, for accreditation targets, the deans of the degree programmes needed valid and reliable outcomes representing the real degree programme level.

## ***Participants***

In sum, 645 participants in five professional bachelor degree programmes for Nursing in the Netherlands were involved in this study: the deans of degree programmes for Nursing in institutions offering professional bachelor of Nursing degrees; ten teachers who were members of educational and examination committees, two members of the central test office; ten students of the final study year, fifteen stakeholders including teachers, the graduates' coordinators, one project manager and 603 responding graduates.

## ***Setting***

The professional bachelor degrees for Nursing in higher education in the Netherlands take four years, and involve a combination of theoretical and workplace-based learning. In this study some specific circumstances were relevant. First, employers signaled the gap between the degrees and developments in the medical sector and recorded this in reports ([Association of Academic Hospitals 2003](#); [Taminiau & Den Boer 2004](#)). Developments with an impact on medical care should have implications for the nursing profession. What are the characteristics of this gap? Second, for most degree programmes it was complicated to respond to the new developments as these require expertise and because a certain dispersion interfered. The fact was that various degree programmes addressed psychiatric and social nursing as well as general nursing. The degree programmes participating in this study focussed mainly on general nursing in hospitals and university medical centres, where the new developments were most pressing. The participating programmes also emphasised a clear distinction with nursing programmes from secondary vocational training.

### ***Procedure and instrumentation***

The procedure of the *ELE* is carried out in the concept of the degree programme level and the steps operationalising and measuring the degree programme level. Each step includes instrumentation to objectify and validate the implementation.

#### ***Step 1: Concept of the degree programme level***

The concept of the degree programme level is elaborated as having a cognitive and an attitudinal component. The cognitive component refers to disciplinary thinking, implying higher order thinking on complex disciplinary problems. The attitudinal component refers to professional attitude, implying the integration of attitudinal aspects of the profession or field the study is addressing. The instruments used were the SOLO taxonomy and the Attitude Development Scheme addressing the expression of the required level of disciplinary thinking and professional attitude in the learning outcomes.

#### ***Step 2: Operationalising the degree programme level***

##### ***Creating themes***

The degree programme level is a concept involving content. The aim of this step is to decide what content is relevant for the level of the degree programme. For this purpose, a study was carried out by experts of participating degree programmes. This required expertise in four areas: the professional and academic fields, external surroundings (referring to specifics of the domain, e.g. a professional organisation, a council of the sector, specific legislation) and the current curriculum. This expertise was provided by three teachers and two external professionals from the field. The experts analysed the four areas on findings and developments of importance for the programme. These were themes substantiated with peer-reviewed literature, for reasons of objectivity and transparency. The themes needed to be consistent with disciplinary thinking and professional attitude and had to be created systematically, in order for them to generate evidence of the measurement's validity (Field 2009, 783). The developed themes were mapped and discussed and resulted in themes.

##### ***Accepting themes***

The themes that were developed by the experts were presented to various stakeholders from the participating degree programmes who were students, teachers and management, mostly united in educational and examination committees. These committees exist in each degree programme in accordance with the Dutch law. The stakeholders indicated the degree of

relevance of the themes for the degree programme level resulting from the previous step on a five-point scale ranging from 'not relevant' to 'very relevant'. It was decided to accept the themes with the highest percentages of relevance. So, here face validity was employed in accordance with the definition of Kane (2006, 36). The aim of this step was to enhance the acceptance of the themes as being valid. The participants now also knew in what direction the evaluation should develop.

### *Developing learning outcomes*

Once the themes were accepted, they were specified into intended learning outcomes, which were statements, written by teachers who were members of educational and examination committees from the student's perspective, reflecting the degree programme level (Biggs & Tang 2007, 70). It was essential to express the required degree of complexity in the formulation of the learning outcomes.

The SOLO taxonomy was used for the learning outcomes expressing the required level of disciplinary thinking. SOLO stands for Structure of the Observed Learning Outcome and provides a systematic way of describing how a learner's performance grows in complexity when mastering many academic tasks (Biggs & Collis 1982). To discriminate between the levels of this taxonomy, consistency is the central notion. This evolves from no consistency in the student's thinking, via no inconsistency into resolving inconsistencies. Biggs and Tang (2007) elaborated the taxonomy for learning at university and presented verbs for intended learning outcomes from the SOLO taxonomy (Table 1).

Insert Table 1 about here

The Attitude Development Scheme was used for the learning outcomes expressing the required level in professional attitude (Table 2). The Scheme provides an approach to indicate how a learner's attitude develops in complexity when incorporating the attitudinal aspects of the profession. The Attitude Development Scheme distinguishes the phases concerning professional and basic attitude.

Insert Table 2 about here

***Validating learning outcomes***

The developed learning outcomes were discussed and finally assessed by other expert teachers, members of examination or curriculum committees, in a test of content validity, resulting in Cohen's Kappa Coefficient. They validated the degree to which the learning outcomes covered the respective themes.

***Step 3: Measuring the degree programme level***

In this step, the measures are mainly concentrated on enhancing validity and preventing bias. The validated learning outcomes were formulated for a questionnaire based on a five-point Likert scale ranging from 'too little' through 'somewhat' to 'more than satisfactory'. The respondents were familiar with the numbers of the scale. Items are generally shorter and more concisely worded than the learning outcomes. Candidate respondents were recent graduates who had completed the programme and had graduated no more than two years previously. They were asked to indicate the extent to which they had mastered the learning outcomes.

The draft questionnaire was pre-tested on ten students in their final study year. It was most important that the candidate should understand immediately the intended kernel of the question. If the student hesitated, this suggested there might be ambiguities. Difficult syntax, mismatching questions and answers, overlapping categories, missing categories and other problems could be traced using coding systems. These pre-tests supported the improvement of the questions' wording and the layout of the web-based questionnaire.

The data collection requires accurate organisation handled by administrators and coordinators of graduates. Manuals were composed to attune the administration procedure and to enhance the awareness of validity of the graduate coordinators from the participating degrees. Data were collected from one central office by graduates who were personally invited by the deans. The respondents' anonymity was guaranteed.

The data were analysed for representativeness, construct validity, and scale reliability.

**Results*****Created themes and their acceptance by stakeholders***

Various stakeholders, i.e., teachers, students and management from the educational and examination committees from the five degree programmes, fully agreed with four of the themes, i.e., nurse-person relationships, nursing diagnosing, nursing care plan, and nursing science.

Nurse-person relationships are considered from the humanising viewpoint of helping the patient and from the dehumanising viewpoint of the business-driven healthcare organisation. In both views, the professional attitude is meaningful. Nursing is a profession that concentrates on assisting human beings in order to help, support, facilitate, and enable them to maintain or recover their well-being in a beneficial manner that makes cultural sense, and helps patients to face disability or death (Gámez 2011). Nurse-person relationships also imply creating a safe setting for the patient to prevent infections and other complications (Pool, Pool-Tromp, Veltman-Van Vugt & Vogel 2004). However, efficiency, technology, and business practices drive healthcare systems. This type of system focuses on outcomes related to efficiency and effectiveness from an economic oriented perspective, instead of the concern for the nurse-person relationship (Milton 2011). Nevertheless, the professional attitude of the nurse is characterised by ‘discretion’, an umbrella term referring to respect for the patient, accurate and confidential handling of the patient’s data, skilful, correct, tactful, appropriate, sensible, and reserved consideration for the patient. The attitudinal aspects are expressed in various actions of the nurse: when nursing the patient, and discussing the patient with family, doctors, and colleagues.

Nursing diagnosing is a term covering standardised actions indicated by scientific terms. The nurse makes a diagnosis by gathering, analysing and validating data. Based on the diagnosis, the nurse decides the interventions. The nursing diagnosis is distinguished from medical diagnosis. The nursing diagnosis refers to the human response aiming to support recovery (Axelsson, Bjövell, Mattiasson & Randers 2006).

The nursing care plan outlines the nursing care to be provided to the patient. The care plan is a set of actions the nurse will implement in response to the nursing diagnosis in support of the medical diagnosis. The plan’s target is to improve the quality of the patient’s life, health and comfort. The plan is based on an understanding of the patient’s responses to the nursing treatment and consideration of the patient’s view of the progress of his/her disease and developments in nursing (Palardy & March 2011).

Nursing science has become an important source for developments in nursing. This discipline is relatively young. The nursing profession has become more complex through the influence of new medicines, new technologies, and new research-based guidelines. The enigma of defining nursing science is preceded by defining nursing, science, research and theory-guided and evidence-based practice (Barrett 2002).

***Developed learning outcomes related to the four themes***

The members of the educational and examination committees decided that the ‘relational’ level is required for the professional bachelor degree programmes for Nursing. On the fourth, relational level of SOLO the student is able to think qualitatively by applying higher order thinking activities without inconsistencies such as critical applying, analyzing, and evaluating. The decision for the fourth level took place after general agreement was achieved through discussions on the explanation of this level (Biggs & Tang 2007, 76-82), analysis of previously used exams and the substantiation of the themes.

The members of the educational and examination committees decided also that the level of ‘organising’, possibly ‘characterising’, is needed for the professional bachelor degree programmes for Nursing. On the fourth level the student is able to integrate aspects of professional attitude without inconsistencies. The cycles of self-regulation are observably to be effective in this level. The decision on the required level took place after general agreement was achieved in consistency with the discussions of disciplinary thinking, the explanation of the attitudinal development (Krathwohl, Bloom & Masia 1974), the cycles of self-regulation (Zimmerman 2006) and the substantiation of the theme referring to the professional attitude.

Some examples of learning outcomes in disciplinary thinking, worded with the support of the SOLO taxonomy, are:

- Explain the relevance of the nursing method for your group of patients
- Analyse laboratory results
- Assess the outcomes of physical examination
- Conclude the validity of your nursing diagnosis

Aspects of the professional attitude are expressed in actions of the nurse, when examining and nursing the patient, in communication with and about the patient with doctors and family.

A few examples, specifying the themes by intended learning outcomes for ‘professional attitude’, worded with the support of the Attitude Development Scheme are:

- Examine the patient physically with respect.
- Quietly clarify the nurse’s intervention to the patient
- Care for the patient according to nursing standards.

### ***Validation of the learning outcomes***

Thirty-one learning outcomes were developed, discussed and finally assessed by expert teachers from other degree programmes in a test of content validity, resulting in Cohen's Kappa Coefficient 0.78, signifying a good level of validity. The validated learning outcomes are represented in Table 4, operationalised by creating and validating themes and specifying these by learning outcomes expressing the level of the professional bachelor degree programme.

### ***Data analysis***

The data were analysed for representativeness, construct validity, and scale reliability.

*Representativeness.* The potential and actually measured numbers of respondents are relevant for the quantitative representativeness: 700 graduates were invited to participate in the survey. The gross number of respondents was 603 (86%). The response rate was comparable across the five degree programmes: it ranged between 114 (83%) and 127 (89%). A total of 133 (19%) respondents were removed: 79 (11%) had graduated  $\geq$  two years ago, and 54 (7%) responded with  $\geq$  30% missing values (they did not answer 9/31 questions). Thus the net number of respondents is 470 (67%). The gross representativeness of 603 is a very good result, and the net representativeness of 470 is more than satisfactory. The real numbers of respondents are sufficient for meaningful analysis of construct validity (Snijkers 2002). The relationship between the measured and real characteristics of the respondents is relevant for the qualitative representativeness. Table 3 is limited to the most important characteristics and the net numbers of respondents. The spread of the variety measured in the five samples agrees approximately with the real variety as demonstrated by labour market surveys of the programmes. The 'Cases with characters A-E' represent the five single degrees. Most respondents had studied the full-time variant, except in Case C. Most respondents in all cases were nurses in a hospital unit in intensive clinical care, and worked in general hospitals and university medical centres (Table 3).

Insert Table 3 about here

*Construct validity.* The full sample of the five degree programmes (N=470) is analysed with Principal Component Analysis (PCA), which is appropriate for exploratory data, assessment, and evaluation of treatments. It is a technique for identifying groups or

clusters of variables and is used to understand the structure of a set of variables, to construct a questionnaire, to measure an underlying variable, and to reduce a dataset to a manageable size while retaining as much of the original information as possible. In the full sample, PCA identified four conceptualised themes and measured four components with eigenvalues  $\geq 1$  and explained 76% variance, which is a good percentage. The result is that the conceptualised themes were affirmed as existing constructs (Creswell 2007, 204) and that from the original learning outcomes eventually twenty-one remained, meaning a reduction of nine items. Sixteen loadings are between .70 and .80 which is good, and five are around .50, which is mediocre (Field 2009, 788). The data of the full sample meet the criteria of construct validity, implying that there is evidence that the content of the test corresponds to the content of the construct that it was designed to cover (Field, 783).

*Reliability.* The components are measured on scale reliability with Cronbach's alpha coefficient. The alpha coefficients are  $\geq 0.81$ , meeting the norm  $\geq 0.70$  for measurements at groups (COTAN 2011) (Table 4). The four components (nursing science, nursing care plan, nurse-person relationships and nursing diagnosing) meet the norm for scale reliability.

### ***Analysis of the components***

The first component, Nursing Science (N=469) explains 47% of the variance and the loadings are from .81, which is very good, as can be seen at the bottom of Table 4. Most of the learning outcomes are the same as those conceptualised. It is noticeable that the scientific terms cluster this component. It is necessary to draw conclusions on the learning outcomes of this component and its setting, which is confirmed by its good scale reliability  $\alpha = .89$ .

The second component refers to the Nursing Care Plan (N=470) which is well-known by the nurses. This component explains 14% variance and includes four excellent loadings from 0.81, and two mediocre ones from .56. These two are probably dated learning outcomes. In sum, four learning outcomes were discarded; three of these were related to the new setting and seem redundant and one was dated. The Care Plan now contains necessary and feasible learning outcomes that are confirmed by good scale reliability  $\alpha 0.84$ .

The third component, Nurse-Person Relationships, refers to the *humanising* and *dehumanising* aspects of the nursing profession (N=468). The human aspects refer to assisting the patient within the dehumanising context of a business-driven organisation. Five learning outcomes explain 7% variance. Three loadings are excellent to good. However, three of the learning outcomes from this component were discarded. Two other loadings are mediocre to low and are related to 'human' aspects: 'examining the patient physically' and

‘using pharmacological knowledge when examining the patient physically’. The latter outcomes can be linked to problems with these learning outcomes, as nurses in other countries encounter problems with these outcomes too (Dilles, Vander Stichele, Van Bortel & Elseviers 2011). The component included good loadings and good scale reliability  $\alpha$  .81.

The fourth and final measured component, Nursing Diagnosing is the most reduced component. Three learning outcomes were discarded; two were measured in other components. The conceptualised theme is not a framed theme for the nurses (N=442). Three learning outcomes explain 7% variance. Two loadings of the remaining learning outcomes are good and the scale reliability  $\alpha$  0.93 is also good.

Insert Table 4 about here

The single samples of the five degree programmes (N=89-99) are also analysed with Principal Component Analysis (PCA). The four themes were measured in components with eigenvalues  $\geq 1$  and explained 72%-80% variance, which are good percentages. The mediocre and low loadings of the full sample are caused by the loadings in the single samples from questions 12 and 13, 17 and 18 and 19, 20. The (dated) questions 12 and 13 appear in the second as well as the fourth component, implying that the learning outcomes ‘indicating to what degree lifestyle influences the patient’s disease’ are affirmed by respondents of degree programmes D and E as part of the Nursing Care Plan (component 2) and interpreted by respondents from degree programmes A, B and C as part of Nursing Diagnosing.

Questions 17 and 18 arise at the second as well as third component, implying that the learning outcomes ‘examine the patient physically’ and ‘integrate pharmacological knowledge when examining the patient’ are understood as part of the Nursing Plan by respondents from programmes B and C and as part of Nurse-person relationships by respondents from degree programmes A, D and E.

Questions 19 and 20 appear at component one as well as four, implying that the learning outcomes ‘analyse the laboratory’s results’ and ‘assess the outcomes of the physical examination’ are understood by respondents of degree programmes D and E as part of the Nursing Science and are interpreted by respondents from degree programmes A, B and C as part of Nursing Diagnosing.

The data of the single samples meet the criteria of construct validity, implying there is evidence that the content of the test corresponds to the content of the construct it was designed to cover (Field, 783).

*Calculated degree programme level.* In deciding to what extent the degree programme level has been achieved, the norm was established at 3.00, meaning results from this point indicate achievement of the degree programme level (Table 5). Based on the valid and reliable outcomes, z-scores were calculated from mean and standard deviation. The components meet the norm for the degree programme level if  $\geq 50\%$  of the respondents scored  $\geq 3.00$ .

Insert Table 5 about here

It is concluded that most problems are in the first component, nursing science. Four outcomes confirm that nursing science is relatively young and that it is not yet present in the Nursing degree programmes. In line with nursing science, there are problems with nursing diagnosing. These components imply that the actions in nursing are more standardised and defined with scientific terms, with the target of establishing accurately the nursing interventions. Most actions of the care plan can be evaluated too and the nursing actions of relationships are intended to refer to attitudinal aspects of discretion.

### **Conclusion and discussion**

The three sub-questions and the main research question ‘*Can the degree programme level in higher education be evaluated validly and reliably by the Educational Level Evaluator?*’ can now be answered:

1) The degree programme level was conceptualised in disciplinary thinking and professional attitude. Disciplinary thinking implies higher order thinking on complex disciplinary problems. Professional attitude implies higher order integration of attitudinal aspects of the profession or field the study is addressing.

2) The Nursing degree programme level was operationalised in four themes. These were analysis-based created in consistency with disciplinary thinking and professional attitude: nurse-person relationships, nursing diagnosing, nursing care plan, nursing science. Various stakeholders, teachers, students and management combined in committees for education and examination of the five degrees accepted these four of the themes, resulting in good face validity.

The accepted themes were specified by student learning outcomes that had to express the required level. Achievement of this goal was supported by two instruments: the SOLO

taxonomy and the Attitude Development Scheme. The learning outcomes were content validated by teachers of the degree programmes.

3) To measure the learning outcomes these were formulated into items for a questionnaire that was pre-tested and adjusted. Recent graduates indicated to what extent they had mastered the required learning outcomes. The data were analysed on validity and reliability. PCA identified the four conceptualised themes as measured components explaining 76% variance. Scale reliability of the components was  $\alpha = \geq 0.81$ , which is a good result. For the single samples, the four measured components explained 72%-80% variance and scale reliabilities  $\alpha \geq 0.78$ . This study demonstrated that the degree programme level can be evaluated validly and reliably with the *ELE*.

The instruments of the SOLO taxonomy and the Attitude Development Scheme should be considered critically. For the taxonomy, we did not use the verbs literally in all cases. A few times we applied more accurate terms to reduce the risk of ambiguity (Chan, Tsui, Chan & Hong 2002) e.g. ‘criticise the reliability of the nursing diagnosis’ instead of ‘review’ and ‘analyse the laboratory results’ instead of ‘review’.

Some outcomes related to ‘discretion’ fell at the pre-test. For example, the following italicized terms rendered too much dispersion: ‘*Quietly* clarify the (nurse’s) intervention to the patient’ and ‘Examine the patient physically with *respect*’. The pre-testers did not make the distinction between the italicized question concerning ‘professional attitude’ and the rest of the question, resulting in bad measuring results.

The norm used for the evaluation of the degree programme level is rather low. The reason for this was that this was the first time that the level has been conceptualised with disciplinary thinking and professional attitude, and the degree programmes did not provide these in all cases.

This work is relevant to the international research community of assessment and evaluation because it presents a method of evaluating validly and reliably the level of degree programmes in higher education. Taking decisions and drawing conclusions about this level requires valid and reliable outcomes because invalid and unreliable outcomes lead to false decisions and conclusions that may have important consequences. This method is also relevant to determine the final attainment level of the higher educational degree programmes. The method lays a solid base for developing and reviewing programmes. This final attainment level frames the programme and is elaborated into intended learning outcomes and assessments for the various phases of the programmes. These learning outcomes facilitate the creation of the activities that will enable the students to achieve the outcomes most

effectively. Then exams can be developed that correspond with the learning outcomes and the effectiveness of the activities (Rexwinkel, Haenen & Pilot in press). Furthermore, this method can create clarity about the degree programme level for various stakeholders, among them students and alumni. It is crucial for them to be certain of the quality of the degree programme level. Based on this information, they make important decisions. Clarity about the degree programme is also relevant for students applying for degree programmes abroad and for receiving higher education institutions, because they also need to know the details of the applying student's degree programme level. The world of the graduate is international. They can apply for a job in other countries, which means the employer needs to know the details of the degree programme level of the applicant. Most higher education institutions in Europe are public institutions that have implemented the bachelor and master structure. They are all obliged to deliver empirical evidence for their degree programme. With this study we wanted to give an empirical underpinning for a method that aims to evaluate the degree programme level validly and reliably.

## References

- American Psychological Association. 2012. *Ethical Principles of Psychologists and Code of Conduct*. <http://www.apa.org/ethics/code/index.aspx#>.
- Association of Academic Hospitals in the Netherlands. 2003. *Competences for the professional bachelor Nurse in University Medical Centers*. Rotterdam: Overbosch Communication.
- Axelsson, L., C. Bjövell, A.-C. Mattiasson, and I. Randers. 2006. Swedish Registered Nurses' incentives to use nursing diagnoses in clinical practice. *Journal of Clinical Nursing* 15(8): 936-946.
- Barrett, E.A. 2002. What is Nursing Science? *Nursing Science Quarterly* 15(1): 51-60.
- Bellingham, L. 2008. Quality assurance and the use of subject level reference points in the UK. *Quality in Higher Education* 14(3): 265-278.
- Biggs, J., and C. Tang. 2007. *Teaching for Quality Learning at University*. 3<sup>rd</sup> ed. Maidenhead, Berkshire: Society for Research into Higher Education & Open University Press.
- Biggs, J.B., and F. Collis. 1982. *Evaluating the quality of learning*. New York: Academic Press.
- Chan, C.C., M. Tsui, M. Chan, and J. Hong. 2002. Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: an empirical study. *Assessment & Evaluation in Higher Education* 27(6): 511-527.
- Cosgrove, R. 2011. Critical thinking in the Oxford tutorial. *Higher Education Research & Development* 30(3): 343-356.
- COTAN [Committee on Test Affairs Netherlands]. 2011) System of assessing the quality of tests of psychologists/NIP. Retrieved from <http://www.psynip.nl>.
- Creswell, J.W. 2007. *Qualitative inquiry & research design*. London: Sage.
- Crochet, M., L. Harvey, H. DuToit, A. Sursock, M. Douterlungne, G. De Fraeye, and C. Hover. *Report of the Committee for the Review of the VLIR Quality Assurance Unit*. European Association of Quality Assurance (ENQA). Retrieved from [http://www.enqa.eu/pubs\\_review.lasso](http://www.enqa.eu/pubs_review.lasso) (2009).
- Dilles, T., R. Vander Stichele, L. Van Bortel, and M. Elseviers. 2011. Nursing students' pharmacological knowledge and calculation skills. *Nurse Education Today* 31: 499-505.

- European Association of Quality Assurance (ENQA). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. 3rd ed. Retrieved from [http://www.enqa.eu/pubs\\_lasso](http://www.enqa.eu/pubs_lasso) (2009).
- Ewell, P. 2010. Twenty years of quality assurance in higher education. *Quality in Higher Education* 16(2): 173-176.
- Field, A. 2009. *Discovering statistics using SPSS*. 3rd ed. London: Sage.
- Gámez, G.G. 2009. The nurse-patient relationship as a caring relationship. *Nursing Science Quarterly* 22(2): 126-127.
- Joint Quality Initiative. *Dublin Descriptors*. Retrieved from: <http://www.jointquality.nl/> (2011).
- Kane, M.Y. 2006. Validation. In *Educational Measurement*. 4<sup>th</sup> ed., ed. R. L. Brennan, 17-64. Westport, CT: Praeger.
- Konrad, H., F. Leynse, M. Crochet, A. Sursock, C. Campbell, D. Dill, S. Neetens, and C. Hover. 2007. *Report of the Committee for the Review of the Accreditation Organisation of the Netherlands and Flanders (NVAO)*. European Association of Quality Assurance (ENQA). [http://www.enqa.eu/pubs\\_review.lasso](http://www.enqa.eu/pubs_review.lasso).
- Krathwohl, D., B. Bloom, and B. Masia. 1974. *Taxonomy of educational objectives*. New York: David McKay.
- Kuh, G.D., J.C. Hayek, R.M. Carini, J.A. Ouimet, R.M. Gonyea, and J. Kennedy. 2001. *NSSE Technical and Norms Report*. Bloomington: Indiana University Centre for Postsecondary Research and Planning.
- Milton, C.L. 2002. Ethical implications for acting faithfully in the nurse-person relationship. *Nursing Science Quarterly* 15(1): 21-24. Retrieved from <http://nsq.sagepub.com/content/15/1/21>.
- Morrow, G., N. Johnson, B. Burford, C. Rothwell, J. Spencer, E. Peile, C. Davies, et al. 2012. Preparedness for practice: The perceptions of medical graduates and clinical teams. *Medical Teacher* 34: 123-135.
- Nederlands-Vlaamse accreditatieorganisatie (NVAO) Assessment frameworks for the higher education accreditation system. *Programme assessment (limited/extensive)*. Retrieved from [http://nvaio.com/page/downloads/DEFINITIEVE\\_KADERS\\_OPLEIDINGSBEOORDLING\\_22\\_november\\_2011\\_English.pdf](http://nvaio.com/page/downloads/DEFINITIEVE_KADERS_OPLEIDINGSBEOORDLING_22_november_2011_English.pdf)

- Nusche, D. 2008. *Assessment of learning outcomes in higher education*. Organisation for Economic Co-operation and Development (OECD), Directorate for Education. Retrieved from <http://www.oecd.org/dataoecd/13/25/40256023.pdf>.
- Palardy, L.G., & A.L. March. 2011. The circle of caring model. *Nursing Science Quarterly*, 24(2): 120-125. <http://nsq.sagepub.com/content/24/2/120>.
- Pool, A., C. Pool-Tromp, F. Veltman-Van Vugt, and S. Vogel. 2001. *With the eye on the future*. Netherlands Institute for Care and Well-being (NIZW). Utrecht: Publishing House NIZW.
- Rexwinkel, G.B., Haenen, J.P.P. & Pilot, A (2012a). Evaluating the level of degree programmes in higher education: conceptual design. *Assessment & Evaluation in Higher Education*. DOI: 10.1080/02602938.2012.755150.
- Rexwinkel, G.B., Haenen, J.P.P. & Pilot, A (2011). Quality Assurance in higher education: analysis of grades for reviewing course levels. *Quality & Quantity*. DOI: 10.1007/s11135-011-9481-6
- Rutowski, L., & D. Rutowski. 2010. Getting it ‘better’: The importance of improving background questionnaires in international large-scale assessments. *Journal of Curriculum Studies* 42: 411-430.
- Sandahl, R., M. Rohlin, M. Waerness, J. Brennan, and K. Sjunesson. *Evaluation of the Swedish National Agency for Higher Education*. European Association of Quality Assurance (ENQA). Retrieved from [http://www.enqa.eu/pubs\\_review.lasso](http://www.enqa.eu/pubs_review.lasso) (2006).
- Snijkers, G.J.M.E. 2002. Cognitive laboratory experiences. Doctoral Dissertation, Utrecht University. [Central Office of Statistics, Netherlands].
- Sternberg, R.J. 2003. What is an “Expert Student?” *Educational Researcher* 32(8): 5-9.
- Taminiau, F., & P. Den Boer. 2004. [*Position of professional bachelor Nurses in general hospitals. Final report*] Sectorfondsen Zorg en Welzijn.
- Zimmerman, B.J. 2006. Development and adaptation of expertise: The role of self-regulatory processes and beliefs. In *The Cambridge handbook of expertise and expert performance*, ed. K Ericsson, N. Charness, P. Feltovich, and R. Hoffman, 705-742. New York: Cambridge University Press.

Table 1

*SOLO taxonomy (Biggs & Collis 1982) with examples of verbs (Biggs & Tang 2007)*

SOLO-level and description	Verbs
5. Extended abstract. There is recognition that the given example is an instance of a more general case. Hypotheses about examples not given are entertained and the conclusions are held open.	Reflect, create, compose, invent, make an original case, theorise, hypothesise, generalise, generate
4. Relational. Most or all of the evidence is accepted, and attempts are made to reconcile. Conflicting data are placed into a system that accounts for the given context. Transitional. There is a hint that closure, or a firm conclusion is not inevitable. There is a suggestion that a relating principle might account for the situation, but this is not spelled out.	Apply, integrate, analyse, explain, conclude, review, argue, transfer, make a plan, debate, construct, solve a problem
3. Multi-structural. Several consistent aspects of the data are selected, but any inconsistencies or conflicts are ignored or discounted so that a firm conclusion is reached. Transitional. Any inconsistencies are noted: Several aspects are recognisable but the student is unable to reconcile them.	Classify, describe, report, discuss, illustrate, select, compute, sequence, outline, separate
2. Uni-structural. An answer is based on only one relevant aspect of the presented evidence so that the conclusion is limited and probably dogmatic. Transitional. An attempt to handle two aspects of the evidence is made, but they may be inconsistent and hence no firm conclusion is reached.	Write, label, identify, recognise, count, define, find, match, memorise, quote
1. Pre-structural. Student avoids the question (denial), repeats the question (tautology), a firm closure based on transduction. Transitional. Student attempts to answer the question but only partially graphs a significant point.	Using tautology to cover lack of understanding, show little evidence of relevant learning.

Table 2

*Attitude development scheme*


---

 Degrees of complexity / level and description
 

---

5. **Characterising.** At the highest level the student is able to place most aspects of the professional attitude consistently into a hierarchical control of his own behaviour. He is able to behave consistently with the aspects of the professional attitude.
  4. **Organising.** The student processes more than one aspect, and can organise these into a system without inconsistencies within the given system. He is able to relate aspects of the professional attitude and integrate these in a system. The self-regulation cycle(s) become more effective.
  3. **Valuing.** The student demonstrates that the required aspects of the professional attitude are becoming observable, and develops to consistency. The student understands the underlying meaning of the attitudinal aspects, although the student's responses are still inconsistent with the attitude he has identified,
  2. **Responding.** He accepts and tends toward the intended behaviour. Above that, he concentrates on the cycle(s) of self-regulation. He orients on goal, task, self-efficacy and outcome expectations to improve his behaviour through self-control and self-instruction. In the self-reflection phase the student determines whether he is mastering an aspect.
  1. **Receiving.** The student becomes aware of the aspects of the professional attitude. He does not yet feel the need for consistency of all the aspects. He also becomes aware of self-regulation cycle(s).
-

Table 3

*Overview of the respondents' characteristics from the five single degree programmes*

Characteristics	Case A N=99	Case B N=94	Case C N=89	Case D N=89	Case E N=99
Year of graduation 2003-2005 <sup>1</sup>	99	94	89	89	99
Study Variant					
Full time	85	46	28	49	61
Part time	10	22	07	16	13
Dual	04	26	54	24	25
Position					
Nursing in a hospital <sup>2</sup>	56	39	59	49	52
Nurse-visitor / Team manager	06	18	07	07	08
Other positions within nursing	25	23	20	26	32
Other positions outside nursing	08	06	02	05	04
Unknown	04	08	01	02	03
Type of Care					
Intensive clinical care	53	42	47	43	47
Chronic / psychiatric nursing	10	13	16	18	16
Various types of care	32	35	25	22	27
Unknown	04	04	01	06	09
Type of Organisation					
General hospital	42	31	35	38	30
University medical centre	11	04	30	14	20
Home care	09	26	08	05	11
Mental health care	10	12	07	18	13
Other healthcare organisations	23	20	09	14	16
Unknown	04	01	--	--	09

<sup>1</sup>Recent graduates <sup>2</sup>Nursing in a hospital's unit e.g. cardiology, oncology, intensive care, anaesthesiology, endoscopy, neonatology

Table 4  
*Summary of Principal Component Analysis results for the level of Nursing Degree Programmes (N=470)*

<i>Item</i>	Rotated Component Matrix			
	Nursing Science	Nursing Care Plan	Nurse-Person Relationships	Nursing Diagnosing
<i>Are you able to:</i>				
01 Explain the relevance of the nursing method for your group of patients	<b>.88</b>	.14	.17	.13
02 Apply treatments within the pattern of nursing interventions	<b>.88</b>	.14	.16	.19
03 Conclude the validity of your nursing diagnosis	<b>.87</b>	.11	.21	.22
04 Criticise the reliability of the nursing diagnosis	<b>.85</b>	.13	.02	.16
05 Consider the nursing guidelines in practice	<b>.82</b>	.16	.05	.11
06 Care for the patient with nursing standards	<b>.82</b>	.24	-.06	.20
07 Make a care plan using measuring instruments	<b>.81</b>	.09	.15	.16
<i>Make a care plan,</i>				
08 with criteria enabling the evaluation of the process as well as the product	.12	<b>.85</b>	.15	.19
09 including the necessary interventions	.09	<b>.85</b>	.20	.07
10 containing new protocols and guidelines	.16	<b>.83</b>	.23	.12
11 concluding the patient's need for nursing care if possible jointly with the patient	.25	<b>.81</b>	.20	-.06
12 indicating to what degree the lifestyle influences the patient's disease	.09	<b>.57</b>	.24	.17
13 concluding to what extent surrounding factors affect the patient's disease	.07	<b>.56</b>	-.01	.19
14 Argue the nursing care of the patient within multidisciplinary deliberation	.15	.27	<b>.85</b>	.13
15 Explain clearly in writing the progress of the patient's disease	.20	.31	<b>.80</b>	.13
16 Clarify the (nurse's) intervention to the patient	.19	.27	<b>.78</b>	.25
17 Examine the patient physically	.19	.13	<b>.50</b>	.21
18 Integrate pharmacological knowledge when examining the patient	.21	.23	<b>.48</b>	.20
19 Analyse the laboratory results	.52	.12	.22	<b>.71</b>
20 Assess the outcomes of the physical examination	.51	.13	.21	<b>.70</b>
21 Apply pathological knowledge when examining the patient	.11	.20	.16	<b>.50</b>
Eigenvalues	8.86	3.07	1.64	1.60
% of variance	<b>46.77</b>	<b>14.18</b>	<b>7.38</b>	<b>7.17</b>
<b><math>\alpha</math></b>	<b>.89</b>	<b>.84</b>	<b>.81</b>	<b>.93</b>

Table 5

*Scales, mean, standard deviation, percentages respondents  $\geq 3.00$  from the five degree programmes and the full sample*

	Case A	Case B	Case C	Case D	Case E	Full sample
<u>Component 1 Nursing science</u>						
M	3.22	2.77	2.69	2.87	2.60	2.85
SD	1.03	0.96	1.05	0.94	0.98	1.01
%	58	41	38	44	34	44
<u>Component 2 Nursing care plan</u>						
M	3.50	3.26	3.14	3.12	2.95	3.20
SD	0.89	0.94	0.94	0.94	0.91	0.91
%	71	61	56	55	48	59
<u>Component 3 Nurse-person relationships</u>						
M	3.44	3.26	3.23	3.12	2.86	3.16
SD	0.88	0.93	1.10	0.87	0.97	1.03
%	69	61	59	56	44	56
<u>Component 4 Nursing diagnosing</u>						
M	3.08	2.93	2.56	2.99	2.51	2.78
SD	1.05	0.96	1.13	1.00	1.06	1.03
%	53	47	35	50	32	42