

Learning outcomes across disciplines and professions: measurement and interpretation

Joakim Caspersenⁱ, Nicoline Frølich, Hilde Karlsen, Per Olaf Aamodt

NIFU Nordic Institute for Studies in Innovation, Research and Education

Abstract

Learning outcomes of higher education are a quality tool in a changing higher education landscape but cannot be seen as neutral measures across professions and disciplines. Survey results from graduates and recent graduates indicate that prevailing measures of learning outcomes yield the same result within and across disciplinary and professional divides. The main interpretation is that learning outcomes must be seen as a valid construct but that the results are highly dependent on the profession and discipline in a way that cannot be reduced to differences in learning outcomes only; measurements of learning outcomes must also be interpreted as mirroring different knowledge structures and knowledge bases in different professions and disciplines. Thus, attempts to make neutral comparisons of learning outcomes between different professions and disciplines are vulnerable to measuring only differences in knowledge structures.

Keywords: Learning outcomes, measurement, professions

Introduction

Learning outcomes of higher education are a key feature in a changing higher education landscape, linked to a range of policy shifts and are applied through a wide range of measurements and approaches. On the international level, higher education learning outcomes (HELOs) are a central part of the Bologna process as well as several European Union and Organisation for Economic Co-operation and Development processes on the development and introductions of qualification frameworks. This attention must be seen in relation to a higher education context seeking approaches for enhanced accountability as part of a wider, major policy shift in public administration, which has re-focused efforts from compliance with rules to production of results (Bovens, 2005; Frølich, 2011). In the contemporary policy context, higher education learning outcomes can therefore be seen as a device for teaching, learning and assessment; but not the least also as a tool linked to governance and management, in the sense that the introduction of HELOs entails a move to a results orientation.

The natural question, following this movement towards outcomes as quality indicator, is how learning outcomes should be measured? In some higher education systems, the use of standardised learning measures (Klein *et al.*, 2005; Klein *et al.*, 2007) has gained ground as a measure of quality in higher education. However, the use of direct measures such as the collegiate learning assessment is far from an international trend and many educational systems rely on more indirect measurements of learning outcomes in their quality assurance systems. Moreover, although the benefits and strengths of direct value-added measures as one component of measuring learning outcomes may be convincing (see e.g. S. Klein, Benjamin, Shavelson, & Bolus, 2007), there are still scepticism towards such measures and a lot of work remains for establishing sound measures of learning outcomes (Green, 2011). Many national policy developments will have to rely on available data sources for measuring learning outcomes, such as self-reported learning outcomes and grades. One example of this development can be found in Norway, where the Norwegian Agency for Quality Assurance in Education (NOKUT), an independent government agency, will introduce self-reported learning outcomes in their new 'barometer' for quality in higher education, to launch this year.

Evidently, learning outcomes are not unambiguous or neutral tools; the principles for assessing learning outcomes, the actual results at individual, study programme and higher education institution level and the dissemination and interpretations of results are generally intended to have a certain impact. In this way, information from assessments may serve as a tool for various stakeholders and actors (those internal and external to higher education systems). Therefore, improving performance indicators and understanding about how learning outcomes can be utilised in performance indicators, is an important undertaking, if HELOs are to play an effective role in the higher education system. One persistent challenge is that the available measures of individual learning outcomes such as tests, grades and self-reported learning outcomes may yield different answers (Hovdhaugen *et al.*, 2007).

Another challenge is how differences in HELOs across disciplines and professions should be interpreted: are similar descriptions of learning outcomes across disciplines measuring the same concept or outcome? Alternatively, do disciplinary differences mean that general descriptions, intended to have validity across disciplines and professions, are measuring different concepts in different contexts?

This article addresses the following question: to what extent do prevailing measures of indirect learning outcomes yield the same result within and across disciplinary and professional divides?

The analyses are done in four different groups, teaching, nursing, engineering and law. Together, these groups provide depth and variety in the comparisons, as they represent different types of hard and soft, pure and applied and general academic and profession-oriented educational programmes. The analyses compare and discuss HELOs in the perspective of three different but interrelated measures: self-reported learning outcomes in the last year of study; learning outcomes measured by grades at graduation; and survey data on self-reported learning outcomes two and a half years after graduation. The purpose is to provide new insights into how learning outcomes can be measured and understood, a question of increasing global importance when assuring the quality in higher education systems.

In the following sections the concept of HELOs is discussed from the perspective of sociology of knowledge. This perspective, emphasising how different disciplines and professions have different social organisations of knowledge. The perspective provides a critical approach to learning outcomes, which is useful before moving on to discussing the empirical measurement of HELOs. Following this, the survey data used in the analyses are presented and the findings reported. The final discussion sums up the findings and discusses whether HELOs can be measured across professions and disciplines with one measure, or whether different professions and disciplines require different measures.

Learning outcomes embedded in disciplinary and professional contexts

One important feature possibly contributing to different learning outcomes across disciplines and professions is that learning outcomes are intrinsically interrelated with acquiring knowledge. They describe individual processes and results but are embedded in disciplinary and professional contexts. The knowledge base of the academic disciplines and professional fields and the norms and values underpinning the socialisation of students are therefore essential aspects of higher education learning outcomes. There are significant differences between hard and soft fields in their degree of paradigm development, specialisation, fragmentation and epistemological beliefs, all of which impact on goals for learning and modes of teaching and instruction (Neumann & Becher, 2002; P. Young, 2010). While teaching in pure academic disciplines relates to a cognitive core, teaching in applied professional programmes is primarily a matter of qualifying for professional practice (Smeby, 2008).

According to Young (2003) it is the social organisation of networks around knowledge that makes for its objectivity and truth. For instance, the system professions use to uphold a high standard in professional practice, through, for example, internal licensing, is a social system organised in order to maintain standards and claims about knowledge, similar to the academic peer-review system. Young's argument is that specialist forms of social organisations such as university disciplines with its social practices are the main guarantees of the objectivity of knowledge. In a similar manner, professions, with their internal certification and control systems, are guarantees for keeping professional practice up to standard. Young described the historical development of disciplines and professions as knowledge-networks, relatively insulated from each other. He argued that knowledge is epistemologically constrained and that knowledge of certain kinds and for certain purposes has to be structured in certain ways.

In this article, learning outcomes are compared with the intent to explore how learning outcomes manifest themselves across different professions and disciplines, when different measures are applied. The groups represent different forms of hard and soft sciences and pure and applied sciences (Muller, 2009), or of general and formative programmes and profession-oriented programmes (Ensor, 2003). Law, according to Ensor, is on the one hand one of the classic academic disciplines and has been institutionalised in universities since the formation

of these institutions, while on the other hand it is also an applied professional programme. Nursing, teaching and engineering are institutionalised in higher education more recently, although at a different pace in different countries. Law was the starting point for the first European university in Bologna at the end of twelfth century (Frijhoff, 1996; Verger, 1992). In the case of Norway, law was one of the first four faculties when the first Norwegian university was established in 1811, together with medicine, theology and philosophy (Slagstad, 2006). Except philosophy, these are applied professional educational disciplines and this, together with the status as an old university faculty and discipline, makes law an interesting hybrid between pure and applied science and between general-formative and applied professional programmes.

Engineering as a higher education field started out with a clearly applied focus, serving the needs of the developing industrial economy and capitalist agriculture in nineteenth century in Europe. Engineering also catered for the specific needs for competence in developing infrastructure in the post-war period (Torstendahl, 1993). Most engineering education was formed out of technical schools who had their background in the labour market and who needed the academic affiliation in their exclusionary strategy for professionalisation through 'academicisation', a strategy that corresponded with the need for developing specialised labour (Torstendahl, 1993). At different institutions across Europe, different importance was given to the practical and scientific mission of what was to become engineering, especially about how the technical aspects of the education should be interpreted. Should 'technical' education mean emphasis on practical knowledge based on scientific knowledge developed elsewhere, or was the separation during education between 'pure' chemistry and the work of the engineer arbitrary (Torstendahl, 1993) During the nineteenth century, a drift from a practical study towards a more abstract programme was found in several institutions. The first programmes started out educating men and managers of private enterprises but developed towards an orientation towards public sector. However, the majority of students continued to complete and start careers in the private sector. In Norway, two different engineering degrees can be awarded: the Master of Science degree has a long standing as the most prestigious technical education and has been a university degree since 1949, with roots back to 1870 and the establishment of the first technical education in 1910 in Trondheim (Slagstad, 2006). The Bachelor of Engineering-degree was formally established in Norway in 2003 but this degree was more-or-less a continuation of the former degree (as also in nursing). Thus, in Norway, two engineering degrees with different educational platforms can be found, with somewhat different academic traditions, where the Master of Science degree can be described as an academic and applied programme, whereas the Bachelor of Engineering is a more general and applied.

Teaching is arguably one of the oldest forms of organised activity and was also institutionalised as a university training programme at an early point in time, often said to be the *École Normale* in Reims in 1685 (giving the name normal schools to teacher educations in many countries (Neather, 1993)). Teaching in Norway is taught as a five-year consecutive master's degree at the universities and as a concurrent bachelor degree in university colleges. Teachers in primary and secondary school hold primarily a bachelor degree (the general teacher education), while teachers in upper secondary hold a master's degree. The predecessors of today's general teacher education in Norway (the programme included in this article) were established in 1826, although some private initiatives had existed since the 18th century (Hagemann, 1992). In 1902, a law was passed, making teacher education a 3-year programme (Karlsen & Kvalbein, 2003) and the programme was extended to four years with the passing of yet another law in 1929. Teacher education has been constantly in tension between two classical curriculum discourses in higher education: the vocational and

professional discourse, oriented towards the application of knowledge, and the discipline-oriented discourse, oriented towards academic standards (Karseth, 2006). In teaching, this tension has manifested itself as a difference between emphasis on discipline orientation in the different subjects taught in schools on the one hand and the practical-didactical elements of teaching on the other. This division is considered to be more distinct in teacher education than other education programmes (Karlsen, 2003).

Nursing started out as an auxiliary position to medicine, held by women. Nursing started its development towards its modern form during the 19th century and has had a rapid development since (Mathisen, 2005). The academic development of nursing started out as a way of improving its position vis-à-vis medicine but also developed as a strategy of distancing nursing from other occupational groups. The double strategy of closure and usurpation (Parkin, 1979) took place through developing nursing as a scientific field in two ways: in the academic tradition of medicine and in a patient and care-oriented bedside-tradition emphasising ethics and reflection. This double strategy has taken different forms in different countries and has led to a diversity and differentiation of the nursing role. However, it seems that nursing is in a dual position with one leg in the medical-academic tradition and another in the professional and applied tradition.

This short presentation of the development of the four groups is useful to point to different kinds of knowledge and skills that are emphasised in different professions and disciplines. With this as a background, one could expect that different educational programmes emphasise different kinds of knowledge and that learning outcomes are strikingly different across groups. The question is whether the differences can be traced empirically in studies of learning outcomes in the four groups. The critical remarks on HELOs, as made by Young (2003) and Ensor (2003), indicate that learning outcomes rarely can be seen as similar across academic fields. However, learning outcomes as implemented in National Qualifications Frameworks have the development of generic skills as a specific outcome, possibly indicating a theoretical and political divide imprinting on the more technical attempts at measuring HELOs. The focus on generic and transferable skills and learning is also evident in attempts to measure learning outcomes, as will be discussed in the next section.

Learning outcomes: measures and methods

Assessing the quality of higher education based on the outcome of learning require reliable and valid test instruments. Perhaps the most commonly used instrument is grades but, as mentioned in the introduction, several other approaches can also be found. Standardised approaches to value added have been developed, and have gained much ground, and students' self-reported learning outcomes when graduating and self-reported learning outcomes after a few years of work-experience are both often-used approaches to assess learning outcomes (H. Karlsen, 2011). In this article, the aim is to examine empirically the relationship between grades and the two kinds of self-reported learning outcomes, that is, indirect measures of learning outcome.

The data used in this article are from two different sources but approaching the same groups. The first data source is self-assessments of competence three years after graduation, taken from the Graduate Survey, administered by the Nordic Institute for Studies in Innovation, Research and Education (Nordic Institute for Studies in Innovation, Research and Education, 2014). The respondents in the survey are encouraged to assess their own competence in several areas and are asked to provide information on their grade-point average upon graduation. By combining this information, it is possible to examine in what areas their

learning outcomes have been greatest and examine the relationship between learning outcomes measured by survey questions and learning outcomes measured by grades.

The second data source is self-reported learning outcomes in the year of graduation, taken from the Norwegian StudData-survey. StudData is a longitudinal panel survey, developed and administered by the Center for Studies of Professions at Oslo and Akershus University College (Center for Studies of Professions at Oslo and Akershus University College, 2014). The database contains information about individuals in a broad range of professional studies and professional careers in Norway. Information on grades in professional education is not available in StudData; learning outcomes are examined by asking the students which types of competence they have acquired because of the education they have followed. Thus, the questions are more directly related to learning outcomes than the questions in the graduate survey. However, as mentioned in the introduction, both are indirect measures of outcomes, in contrast to the more direct approaches found in for example in the collegiate learning assessment (Klein *et al.*, 2007; Klein *et al.*, 2005).

The three measures provide a contrast between formal assessment (grades) and self-assessment. Moreover, by using data from after graduation, after the knowledge acquired is 'tested' on the realities of work, an employability perspective is also included. The analyses are based on cross-sectional information from students who completed their professional education in 2007 (StudData) and 2008 (the Graduate Survey). Table 1 present the different data sources.

Table 1 Overview of data sources

Graduates from 2008: Surveyed 3 years after graduation (Graduate survey)		Graduates from 2007 (StudData).	
<i>Study programme</i>	<i>Type of indicator</i>	<i>Study programme</i>	<i>Type of indicator</i>
Teaching (4 yrs)	Self-reported outcome & self-reported grades	Teaching (4 yrs)	Self-reported outcome
Nursing BA	Self-reported outcome & self-reported grades	Nursing BA	Self-reported outcome
Engineering BA	Self-reported outcome & self-reported grades	Engineering BA	Self-reported outcome
Engineering MA	Self-reported outcome & self-reported grades		
Law MA	Self-reported outcome & self-reported grades		

Engineers with a master's degree are the largest group in the Graduate Survey (470 respondents, 25.4%), while Law students are the smallest group (219 respondents, 11.9%) (Table 2). Nurses are by far the largest group in the StudData-survey (1001 respondents, 42.2%). Law and engineering (MA) are not included in StudData (Table 2). The response rate in StudData was 71% for teachers and nurses and 54% for engineers. In the Graduate Survey the response rate was 44% for teachers, 39% for nurses, 42% for engineers (BA), 51% for engineers (MA) and 48% for law (MA).

Table 2: Educational groups in Kandidatundersøkelsen and StudData

Educational group	Graduate Survey		StudData	
	Freq.	Percent	Freq.	Percent
Law (MA)	219	11.9		
Engineer (MA)	470	25.4		
Engineer (BA)	331	17.9	541	22.8
Nursing (BA)	447	24.2	1001	42.2
Teacher (4 years)	381	20.6	832	35.0
Total	1848	100.0	2374	100.0

As the two data-sources are not matched but directed towards different students, it is not a goal to analyse the relation between responses in StudData and in the Graduate Survey. However, as the surveys are at the same time similar but also clearly different, they provide an opportunity for an exploration of how learning outcomes manifest themselves across different data sources, where the difference in time for gathering responses introduces different kinds of ‘noise’. First, they have different, although similar, wording and variables. Second, the difference in time for gathering responses introduces different kinds of ‘noise’. Therefore, instead of focusing on the exact differences in means between groups or surveys, the interesting outcome is whether patterns are found between different types of measures across groups. If the two different surveys produce the same pattern between the groups, it supports the assumption that group differences affect the results rather than the actual learning outcomes of the groups.

As the analyses are explorative, specific hypotheses are not put forward. Instead, different statistical techniques are applied to examine the relationships between the different measures, in a consecutive order. The information from the surveys are first analysed using exploratory factor analysis (Kim & Mueller, 1978), in order to examine whether different items can be grouped together in different latent constructs, each measuring specific parts of learning outcomes. The number of factors was decided using scree-plot, (a graphical interpretation of the number of factors to extract), which in this case also corresponded to Kaiser’s criterion of excluding all components with eigenvalue less than 1 (Kim & Mueller, 1978). Two criteria for choosing which variables to include in the factors were applied. First, the variable must have a factor loading of more than 0.3 and, second, not decrease the internal consistency of the variable, as measured by Cronbach’s alpha (reported in the last row). Thus, the variable ‘ability to write reports, notes or documents’ is included in the factor ‘Knowledge’, as the inclusion of it in the factor ‘Reflective’ decreases the internal consistency, despite its higher loading on this factor.

The choice of exploratory factor analysis instead of confirmatory factor analyses may seem surprising, given that the data to some extent builds on previous research and theoretic development. However, as discussed in Caspersen *et al.* (2011) the measures used are not identical to surveys such as, for example, the National Survey of Study Engagement (2014) or the REFLEX (2014) study although inspiration and some items are taken from these and other surveys. Thus, it seems inappropriate to approach them in a confirmatory manner, although this approach should be preferred in future research.

Following the factor analyses, the relationship between self-reported grades and self-reported learning outcomes in the Graduate Survey is presented using Spearman’s rank correlation coefficient (ρ). Finally, in order to visualise the patterns of relations between groups, mean scores are presented graphically. The mean scores were estimated on each latent construct, by using multivariate analysis of variance (MANOVA) to identify differences between groups on the mean scores on a linear component of the dependent variables. The results from the MANOVAs are not presented in detail here but only used for understanding the pattern between the groups when using different measures.

Self-reported learning outcomes two to three years after graduation

The first analysis explores the factor structure of the two learning outcomes batteries, in order to compare similarities and differences in factor structure. A factor analysis of the graduates’ assessment of own competence in the Graduate Survey is presented in Table 3.

Table 3: Factor analysis of learning outcomes/self-assessed competence in the Graduate Survey

How would you rate your level of competence on the following measures?	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
<i>I am able to .../I have....</i>	Practical learning outcome	Leadership learning outcomes	Reflective learning outcomes	Knowledge learning outcomes	
coordinate activities	0.72	0.05	0.00	0.03	0.40
exercise authority/leadership	0.23	0.65	-0.03	-0.07	0.43
ethical judgment	0.02	-0.03	0.75	0.00	0.46
reflect upon own practices	0.07	0.06	0.69	-0.06	0.47
mobilise other peoples resources	0.39	0.38	0.17	-0.16	0.47
efficient	0.75	-0.12	0.07	0.00	0.48
work under pressure	0.65	0.06	-0.09	0.16	0.48
work productively in a team	0.59	0.00	0.20	-0.01	0.49
alert to new possibilities	0.51	0.23	-0.05	0.12	0.50
make myself understood	0.27	0.35	0.20	0.02	0.51
quickly acquire new skills	0.28	-0.08	-0.06	0.62	0.52
come up with new ideas and solutions	-0.01	0.39	0.09	0.33	0.53
communicate with customers/clients (-)	0.13	0.27	0.43	-0.05	0.53
present products, ideas and reports	-0.10	0.54	0.05	0.24	0.54
question own/others' ideas	-0.05	0.40	0.15	0.28	0.55
negotiating skills	0.31	0.49	-0.07	-0.02	0.56
analytical thinking	-0.06	0.05	-0.09	0.70	0.56
mastering own professional field	0.19	0.05	0.00	0.46	0.64
write reports, notes and documents	0.08	-0.06	0.34	0.32	0.67
knowledge on other professional fields	0.07	0.18	-0.09	0.44	0.70
talk and write in a foreign language	-0.11	0.04	0.07	0.51	0.72
use a computer and the Internet	-0.02	-0.10	0.15	0.48	0.74
Cronbach's alpha	0.85	0.83	0.77	0.78	

Note: (-) = variable not included in any of the latent constructs. Variable values from 1 to 7, where 1 = 'very low' and 7 = 'very high.' Bold font indicates the single variables included in the factor. Factor rotations at four factors, principal factor analyses with oblique rotation

The first factor found can be labelled 'Practical learning outcomes' and consists of 'ability to coordinate activities', 'ability to mobilise other people's resources', 'being able to plan and organise within strict time frames', 'ability to perform under pressure', 'being able to function as part of a team' and 'being alert to new opportunities that open up'.

The second can be labelled 'Leadership learning outcomes' and consists of 'ability to lead with authority', 'being able to come up with new ideas and solutions', 'being able to present products, ideas or reports', 'being willing to question own and other people's ideas' and 'negotiating skills'.

The third factor is labelled 'Reflective learning outcomes' and consists of the two variables 'ethical reflection skills' and 'ability to reflect upon own practice'.

The fourth and final factor is called 'Knowledge learning outcomes' and consists of 'ability to acquire new knowledge quickly', 'analytical thinking', 'master one's own professional field', 'being able to write reports, notes and documents', 'knowledge on other professional fields', 'being able to talk and write in a foreign language' and 'being able to use PC and the Internet'.

The variables in StudData are presented in Table 4. Cronbach's Alpha for each factor is shown at the bottom of the table. The same criteria are used for factor extraction as was used for the Graduate Survey.

Table 4: Learning outcomes in StudData

Did you acquire these types of competence as a result of your education?	Factor 1	Factor 2	Factor 3	Uniqueness
<i>I am able to .../ I have.../</i>	<i>Social and ethical</i>	<i>Leadership learning outcomes</i>	<i>Practical learning</i>	
broad, general knowledge	-0.13	0.20	0.46	0.72
work-specific knowledge	0.06	0.06	0.44	0.74
knowledge on planning and organising	-0.11	0.08	0.62	0.62
insight into rules and regulations	0.11	-0.15	0.51	0.75
being able to critically reflect upon and assess own work	0.23	-0.01	0.50	0.59
come up with new ideas and solutions	0.00	0.17	0.59	0.53
work under pressure	-0.10	0.41	0.41	0.58
practical skills	0.17	0.08	0.55	0.52
being able to work independently	-0.03	0.48	0.30	0.57
interpersonal skills	0.10	0.42	0.18	0.64
take initiative	0.31	0.38	0.13	0.52
verbal communication skills	0.46	0.34	-0.03	0.53
written communication skills	0.28	0.48	-0.08	0.60
tolerance, ability to appreciate different points of view	0.53	0.28	-0.06	0.52
leadership abilities	0.14	0.43	0.18	0.59
take responsibility and make decisions	0.24	0.51	0.11	0.46
ethical judgement	0.72	-0.03	0.07	0.45
empathy	0.85	0.05	-0.10	0.30
theoretical knowledge(—)	-0.03	0.44	-0.04	0.84
values and attitudes	0.69	-0.09	0.20	0.43
cope with the emotional challenges in my work	0.59	0.15	0.04	0.50
Cronbach's alpha	0.87	0.83	0.81	

Note: (-) = variable not included in any of the latent constructs; * = variable included in more than one latent construct. Variable values from 1 to 5, where 1 = 'not at all' and 5 = 'to a high degree.' Bold font indicates the single variables included in the factor. Factor rotations at three factors; principal factor analysis with oblique rotation.

In the factor analysis of StudData, three factors emerged, in contrast to the four factors found in the Graduate Survey. The first factor is called 'social and ethical' learning outcomes'. The factor consists of the variables 'verbal communication skills', 'tolerance and ability to value other's point of view', 'ability to ethical judgment', 'being able to being empathic with other people', 'values and attitudes' and 'being able to cope with the emotional challenges in the work'.

The second factor is called 'leadership learning outcomes' and consists of the variables 'being able to work under pressure', 'being able to work independently', 'being able to take initiative', 'written communication skills', 'leadership ability', 'being able to take responsibility and make decisions'.

The third and final factor is called 'practical learning outcomes' and consists of 'broad, general knowledge', 'work-specific knowledge', 'knowledge on planning and organising', 'knowledge on rules and regulations', 'being able to critical reflect upon and evaluate own

work', 'being able to think in new patterns', 'ability to work under pressure' and 'practical skills'.

The three factors that emerge are similar to three of the factors found in the Graduate Survey, although somewhat different labels are used to present the nuanced differences. The factor that is not replicated in this survey is the 'knowledge' factor. However, in the StudData survey, the single item of learning outcomes on theoretical knowledge is not included in any of the three factors. The exclusion is based on the criteria for definition of factors, as the inclusion of theoretical knowledge in the 'leadership learning outcomes' dimensions would weaken the reliability of the index, measured by Chronbach's alpha.

Seen together, the two surveys then produce similar but not identical, factors (which is natural, since the wording and items are different in the two and they are administrated at different points in the respondents' professional careers). In the following section, the patterns of learning outcomes across the professions in the two different surveys are examined.

The patterns and differences

In order to compare how different measures of learning outcomes produce similar or different patterns across groups, the mean scores and the standardised mean scores (mean=0, standard deviation=1) on learning outcomes for all groups on the different factors found are compared (Figure 1). Potential group differences were examined using one-way MANOVA. For both surveys, the correlation between the four dimensions ranged from 0.4-0.7. As discussed, the factors found in the Graduate Survey and the StudData survey are not identical but somewhat similar dimensions are found in the two and they can thus be compared.

Figure 1 about here.

The unstandardised means in Figure 1 show that the mean scores are very different across the two surveys, with the Graduate Survey producing mean scores around 5 and the StudData survey around 3.5. The only large difference between the two surveys is found among teachers' reports on leadership skills. In the StudData survey (at graduation), teachers report slightly lower mean (3.5) score than nurses (3.6) and slightly higher than engineers (3.4). In the Graduate Survey, however, teachers have a higher mean score (5.1) than all the other groups (4.8-4.9).

A question that naturally raises itself is why the absolute differences in mean scores between the two surveys are so large. First, the differences in scales used (1-7 *versus*. 1-5) makes it natural that the mean scores are higher in the Graduate Survey. Furthermore, one could also argue that the Graduate Survey, being sent out two and a half years after graduation, includes learning outcome from work as well as from the respondents' studies. This would mean that the difference between the mean scores indicate a 'value added' from graduation to the time of the Graduate Survey. If so, the distinct increase in teachers' leadership skills would imply that they have had a large learning outcome on this dimension after graduation.

To compare the pattern between the two surveys more easily, the standardised means are included in Figure 1. This makes it harder to assess the potential value added but accentuates the similarities in the patterns between the two groups.

Learning outcomes and grades

So far, the differences in learning outcomes as reported in the 22 questions used in the Graduate Survey have been presented. The graduates are also asked to report their grade-point average from their education. This makes it possible to examine the relationship between

learning outcomes as measured by grades and learning outcomes measured by other means. Often, grades are assumed to measure specific competence, in contrast to generic competence, as discussed in previous sections. However, as the analyses do not have grades for different subjects, the grade-point average rather represents overall competence, both generic and specific. Students most frequently report a grade point average of C or B (Figure 2). Very few students admit to a grade point average below C.

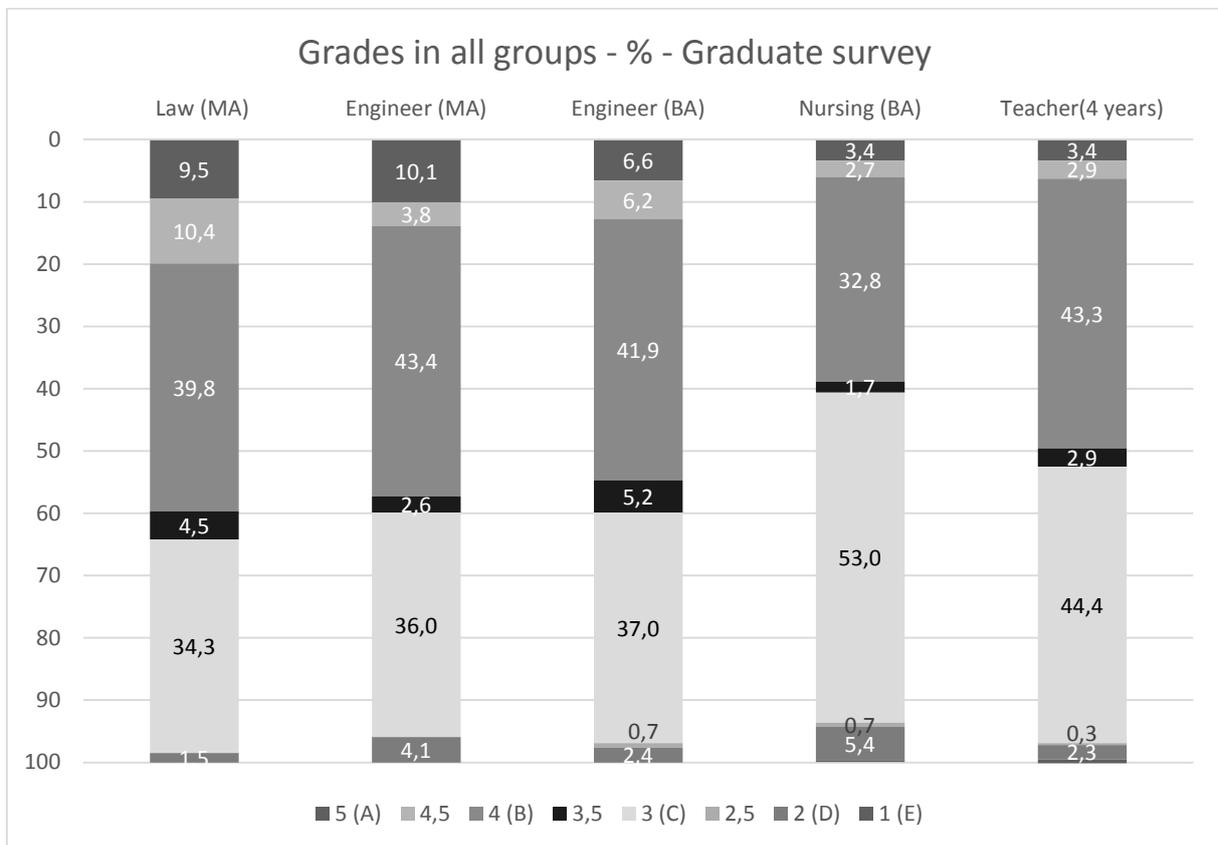


Figure 2: Grades, self-reported two years after graduation

If the share of graduates reporting a grade below average, i.e. C or lower, are compared across groups 35.8% of the law graduates, 40.1% of the two engineer groups, 44.4% of the teachers and 59.3% of the nurses report grades lower than C. Thus, nurses report lower grades than the other groups. At the top of the scale, the two groups of graduates with a MA, law and engineering, have a higher share of students who report an A, 9.5% and 10.1%, respectively, while in the other engineer group 6.6% report an A as grade-point average. Among teachers and nurses, only 3.4% of the graduates report and A as a grade-point average. Thus, nurses seem to be the group with lowest reported learning outcome measured by grade-point average, followed by teachers, while law graduates have the highest reported learning outcome. This is parallel to the learning outcomes as measured by the ‘Knowledge factor’ (Table 4) and supports the interpretation that learning outcomes measured across professions and disciplines differ systematically.

It should be noted that compared to official statistics of the graduates from the same semester and education, the self-reported grades have a higher average. In official statistics, 33–39 % have a grade average lower than C, including failed students (F). There are many potential sources for this difference: most importantly, F students are less likely to be working in the profession, as a passed examination is the entrance qualification, and this is probably the case for D and E students, who are less attractive on the labour market. Furthermore, retrospective

questions can produce biased responses and so could the respondents' tendency towards positive self-representation. However, as previous studies have shown few differences between respondents and population (Arnesen, 2012) it could be argued that the self-reported grades are valid measures.

Table 5: Spearman's ρ between the grade-point average and the four learning outcomes (from Table 4) (latent constructs)

Factors	All five educations	Law (MA)	Engineer (MA)	Engineer (BA)	Nursing (BA)	Teaching (4 years)
Reflective	-0.043	-0.051	-0.024	-0.051	0.032	0.018
Leadership	-0.046	0.0099	-0.026	-0.148*	-0.039	0.043
Practical	-0.074**	-0.009	-0.087	-0.117*	-0.043	0.049
Knowledge	0.133***	0.212**	0.126*	0.036	0.104*	0.104
N	1659	200	415	288	409	347

Note: casewise deletion

For all groups combined, a modest, positive, significant correlation between knowledge learning outcomes and grade-point average and a weak but significant negative correlation between practical learning outcomes and grade-point average is found. (Table 5) When the groups are examined one by one, a positive correlation between knowledge learning outcomes and grade-point average can be found among law, engineering (MA) and nursing graduates and strongest among the law graduates. The negative correlation between practical learning outcomes and grade-point average can only be found among engineering graduates with a BA, however. In this group, a negative significant correlation between leadership learning outcomes and grade-point average can also be found, indicating that higher grade-point average is correlated with less leadership-learning outcome and *vice versa*. For none of the groups can a significant correlation be found between reflective learning outcomes and grade-point average.

Learning outcomes and knowledge structures

The main question in this article was to what extent prevailing measures of learning outcomes yield the same result within and across disciplinary divisions. Based on the analyses presented in Figure 1, it could be argued that stability both within and across groups was observed. The question is whether the stability in patterns is due to the validity of the two different measures used. Alternatively, are the group patterns and relative positions of the groups related to the knowledge structures of the disciplines and professions examined? A case can be made for both arguments, as these two dimensions are intertwined in a way that makes it difficult to distinguish between them. This will be elaborated in the following sections.

One can discuss whether the history and development of the knowledge structures in each of the disciplines and professions is as idiosyncratic as Young (2003) and Ensor (2003) seem to argue. For instance, the professions are, and have always been, a blend of different other disciplines and approaches, with the aim being a competent professional practice. One argument in this article is that the differences in knowledge structures between disciplines make the measurement of learning outcomes challenging.

Learning outcomes, as they are approached in the surveys analysed (which again are inspired by many other often-used measurements of learning outcomes) can best be understood as approaching 'transferable skills'; that is, the kind of general learning outcomes that are said to

be particularly relevant for candidates' employability on today's dynamic and knowledge intensive labour market (sources). The importance of an education that includes transferable skills, not just purely academic learning, is emphasised internationally as well as nationally (NOU, 2003).

Ensor (2003) and Young (2002; 2003) argued critically against the introduction of this kind of learning outcomes, claiming that they create new, arbitrary knowledge fields that cut across academic disciplines and occupational fields. This gives priority to cross-sectoral-level descriptors and standard-based qualifications that are not founded in epistemological arguments about the nature of knowledge. Furthermore, the introduction does not take into account what Ensor referred to as vertical and horizontal differences between types of knowledge, that is, between and within disciplines, and that these rest on fundamental differences. In the analyses, four different academic and professional groups have been examined, which represent different forms of hard and soft sciences and pure and applied sciences, or of general and formative programmes and profession-oriented programmes, with a varied history of institutionalisation in higher education

Engineering is taught based on two, somehow different, educational platforms in Norway, which differ in their academic orientation, although they both have an applied focus. These differences manifest themselves on the knowledge dimension of the learning outcome concept; engineers at bachelor level report lower knowledge than engineers at master's level. Engineers at master's level report also higher learning outcomes on the knowledge dimension than law candidates. Engineers at BA level also report lower knowledge than nurses and teachers. Regarding practical learning outcomes, engineers at both levels report lower learning outcomes than nurses and teachers. They also report lower leadership learning outcomes than teachers. Finally, engineers report lower reflective learning outcomes than nurses and teachers.

Nursing has a history of strong development of its academic status but is, in many respects, an applied and soft profession. Nurses report significant higher practical learning outcomes than law and engineering; higher reflective and social and ethical learning outcomes than engineers and teachers; higher knowledge than teachers and engineers at BA level but lower knowledge learning outcomes than engineers at master's level.

Teacher education has historically found itself ridden by tensions between vocational and professional and discipline orientation. Teachers in the sample report higher practical learning outcomes than engineers, higher leadership learning outcomes than these as well as lawyers, however lower knowledge orientation than engineers at master's level. Teachers report, however, lower social and ethical, reflective and knowledge-learning outcomes than nurses.

Overall, the group differences can very well reflect the knowledge structures and development of the disciplinary fields.

The validity of and relationship between measures

By comparing the measures of learning outcomes in StudData and the graduate survey, similar but not identical dimensions across the different self-reported learning outcomes measures are found. In the Graduate Survey, the four dimensions practical learning outcomes, leadership learning outcomes, reflective learning outcomes and knowledge related learning outcomes, were found. In StudData, the dimensions social and ethical learning outcomes, leadership learning outcomes and practical learning outcomes and the single variable measuring outcome of theoretical knowledge, were found. The items in the social and ethical learning outcomes are quite similar to the reflective learning outcomes found in the graduate

survey. Except for the knowledge-related learning outcomes dimensions, the same dimensions are visible across different operationalisations and surveys.

Thus, the two different measures produce similar dimensions across the two samples, strengthening the argument that practical, knowledge-related, leadership-related and reflective learning outcomes are valid dimensions of the broader concept of learning outcomes; also across professions and disciplines. This argument is further supported by the fact that one of the samples is from graduation and the other is two and a half years after graduation. However, it is necessary in the future to develop, test and establish more unified approaches for measuring learning outcomes in higher education.

Despite this correspondence in dimensions across two different surveys, the low and unexpected correlation with grades in the analyses also questions the validity. Judged only by grade point average, nurses seem to be the group with lowest reported learning outcome, followed by teachers, while law graduates have the highest reported learning outcome. The group differences are more distinct when looking at grades than when looking at the self-reported learning outcomes.

However, more interestingly, the analyses tested how grades vary with the content of learning outcome. *A priori* one would perhaps expect a positive correlation between all learning outcomes and grades, or a neutral relation. The only systematic relation was found between grades and knowledge-related learning outcomes, where the two academic master programmes, law and engineering, have a significant (but not very strong) correlation with knowledge-related learning outcomes. A significant (but even weaker) correlation is also found in nursing, which perhaps can be related to the strong academic orientation found in nursing and the use of knowledge and development of a professional knowledge as a tool for raising the status of the nursing profession. Research (also using StudData) has also shown Norwegian nursing students to be highly oriented towards academic knowledge when entering and emphasising this even more at graduation (Heggen, 2008).

The weak relation between grades and other types of measures of learning outcomes is even further illustrated by the negative relation between grades and leadership learning outcomes and practical learning outcomes among those with a bachelor in engineering. This again highlights the difficulties and challenges in using the same measures of learning outcomes across professions and disciplines as they relate positively to one kind of measures, neutrally to others and even negatively to some; and it all depends on which groups are investigated. Grades, perhaps the most widely used proxy for learning outcomes, do not have a particularly strong relation with learning outcomes in higher education measured by alternative methods.

Grades and the knowledge dimension of learning outcomes are somewhat positively related. At the same time the knowledge dimension of learning outcomes vary among the case groups. As discussed, engineers at master's level report higher knowledge learning outcomes than lawyers, engineers at BA level, nurses and teachers. Based on these results, one could argue that grades reflect better learning outcomes of engineers at graduate level than the other professions. One implication would be that not only do grades just partly capture learning outcomes; they also capture them less in some professions than others. If this is the case, then grades as measures of learning outcome is only partly comparable across disciplines and professions. Similar arguments are also proposed by Bowman, (2010, 2011) who argued that there is hardly any correlation between grades after one year of study and grades at the end of studies and self-reported learning outcomes at the end of study. The data indicates that there is an ambiguous relationship between self-reported learning outcomes and grades.

Conclusion

Based on the explorations of learning outcomes across different measures and professions, it is argued that the data supports the idea that measures of learning outcomes vary systematically between disciplines and professions but also that valid measures can be found. However, the fact that the same or similar dimensions are replicated across different survey instruments does not support the use of any particular instrument. Given the differences in the instruments, it seems reasonable to argue that the similarities in patterns between groups should be interpreted as expressions of differences in knowledge structures, not only as differences in learning outcomes.

Significant and systematic differences across the cases both in strengths of learning outcomes and their content were found. The finding supports the argument that differences to some extent are based on the social organisation of knowledge networks, not only 'real' differences. Thus, the analyses clearly indicate that learning outcomes are ambiguous and multifaceted and illustrate some of the challenges with disentangling learning outcomes from the construction of the discipline or profession.

References

- Arnesen, C. Å., 2012, *Profesjonsutdannedes overgang fra utdanning til arbeid: Arbeidsmarkedstilpasning, yrkesverdier og tilfredshet med arbeid og utdanning*. NIFU-report 13/2012 (Oslo, NIFU).
- Bovens, M., 2005, 'Public accountability'. In E. Ferlie, L. E. j. Lynn & C. Pollitt (Eds.) *The Oxford Handbook of Public Management* (Oxford, Oxford University Press).
- Bowman, N. A., 2010, 'Can 1st-year college students accurately report their learning and development?' *American Educational Research Journal*, 47(2), pp. 466–96.
- Bowman, N. A., 2011, 'Validity of college self-reported gains at diverse institutions'. *Educational Researcher*, 40(1), pp. 22–24.
- Caspersen, J., de Lange, T., Prøitz, T. S., Solbrekke, T. D., & Stensaker, B., 2011, *Learning About Quality – Perspectives on learning outcomes and their operationalisation and measurement* (Oslo, Department of Educational Research, University of Oslo).
- Center for the Study of Professions, 2014, *Presentation of the StudData project*. Available at: <http://www.hioa.no/eng/Research-and-Development/Our-research/FoU-SPS/prosjekter/StudData> (accessed 27 November 2013).
- Ensor, P., 2003, 'The National Qualifications Framework and higher education in South Africa: some epistemological issues', *Journal of Education & Work*, 16(3), pp. 325–46.
- Frijhoff, W., 1996, 'Themes', in Ridder-Symoens, H.D. (Ed.), *A History of the University in Europe*, volume. 2 (Cambridge, Cambridge University Press).
- Frølich, N., 2011, 'Multi-layered accountability. Performance funding of universities'. *Public Administration*, 89(3), pp. 840–59.
- Hagemann, G., 1992, *Skolefolk: Lærernes historie i Norge* (Oslo, Ad notam Gyldendal).
- Heggen, K., 2008, 'Social workers, teachers and nurses from college to professional work', *Journal of Education and Work*, 21(3), pp. 217–31.
- Karlsen, G., & Kvalbein, I.A., 2003, *Norsk lærerutdanning: søkelys på allmennlærerutdanningen i et reformperspektiv* (Oslo, Universitetsforlaget).
- Karlsen, G.E., 2003, 'Ny allmennlærerutdanning - struktur, innhold og kompetanse', in Karlsen, G.E. & Kvalbein, I.A. (Eds.), *Norsk lærerutdanning. Søkelys på allmennlærerutdanningen i et reformperspektiv* (Oslo, Universitetsforlaget).
- Karlsen, H., 2011, *Klare for arbeidslivet? En drøfting av metodiske utfordringer for måling av læringsutbytte i høyere utdanning*. NIFU-rapport 42/2011 (Oslo, NIFU).
- Karseth, B., 2006, 'Curriculum restructuring in higher education after the Bologna process: a new pedagogic regime', *Revista Española de Educación Comparada*(12), 255–84.
- Kim, J.-O., & Mueller, C.W., 1978, *Factor Analysis: Statistical methods and practical issues* (Beverly Hills, Sage).
- Klein, S., Benjamin, R., Shavelson, R., & Bolus, R., 2007, 'The collegiate learning assessment—facts and fantasies'. *Evaluation Review*, 31(5), pp. 415–39.
- Klein, S.P., Kuh, G.D., Chun, M., Hamilton, L. & Shavelson, R., 2005, 'An approach to measuring cognitive outcomes across higher education institutions', *Research in Higher Education*, 46(3), pp. 251–76.
- Mathisen, J., 2005, *Sykepleiehistorie*, Second edition (Oslo, Gyldendal akademisk).
- Muller, J., 2009, 'Forms of knowledge and curriculum coherence', *Journal of Education and Work*, 22(3), pp. 205–226.
- National Survey of Study Engagement, 2014, *Presentation*. Available at: <http://nsse.iub.edu/html/about.cfm> (accessed 3 February 2014).

- Neather, E.J., 1993, 'Teacher education and the role of the university: European perspectives'. *Research Papers in Education*, 8(1), pp. 33–46.
- Neumann, R., & Becher, T., 2002, 'Teaching and learning in their disciplinary contexts: a conceptual analysis'. *Studies in Higher Education*, 27(4), 405–17.
- Nordic Institute for Studies in Innovation, Research and Education (NIFU), 2014, *Presentation of the Graduate Survey*. Available at: <http://www.nifu.no/Norway/Sider/STATISTIKK/Kandidatunders%C3%B8kelsen/Halvtaarsundersokelser/Halvtaarsundersokelser.aspx?ItemId=2715&ListId=3074294f-5cf7-4988-b2da-72ceaa435ff1> [Norwegian only] (accessed 27 November 2013).
- Parkin, F., 1979, *Marxism and Class Theory: A bourgeois critique*. (New York, Columbia University Press).
- REFLEX, 2014, *Info*. Available at: <http://www.fdewb.unimaas.nl/roa/reflex/> (accessed 3 February 2014).
- Slagstad, R., 2006, *Kunnskapens hus*, Second edition (Oslo, Pax).
- Smeby, J.-C., 2008, 'Profesjon og utdanning', in Molander, A. & Terum, L.I. (Eds.), *Profesjonsstudier* (Oslo, Universitetsforlaget).
- Støren, L.A., & Aamodt, P.O., 2010, 'The quality of higher education and employability of graduates', *Quality in Higher Education*, 16(3), pp. 297–313.
- Torstendahl, R., 1993, 'The transformation of professional education in the nineteenth century', in Rothblatt, S. & Wittrock, B. (Eds.), *The European and American University Since 1800. Historical and sociological essays* (Cambridge, Cambridge University Press).
- Verger, J., 1992, 'Patterns', in Ridder-Symoens, H.D. (Ed.), *A History of the University in Europe* (Cambridge, Cambridge University Press).
- Young, M., 2002, 'Contrasting approaches to qualifications', in Evans, K., Hodkinson, P. & Unwin, L. (Eds.), *Working to Learn* (London, Kogan Page).
- Young, M.F.D., 2003, 'National Qualifications Frameworks as a global phenomenon: a comparative perspective', *Journal of Education and Work*, 16(3), pp. 223–37.
- Young, P., 2010, 'Generic or discipline-specific? An exploration of the significance of discipline-specific issues in researching and developing teaching and learning in higher education', *Innovations in Education and Teaching International*, 47(1), pp. 115–24.

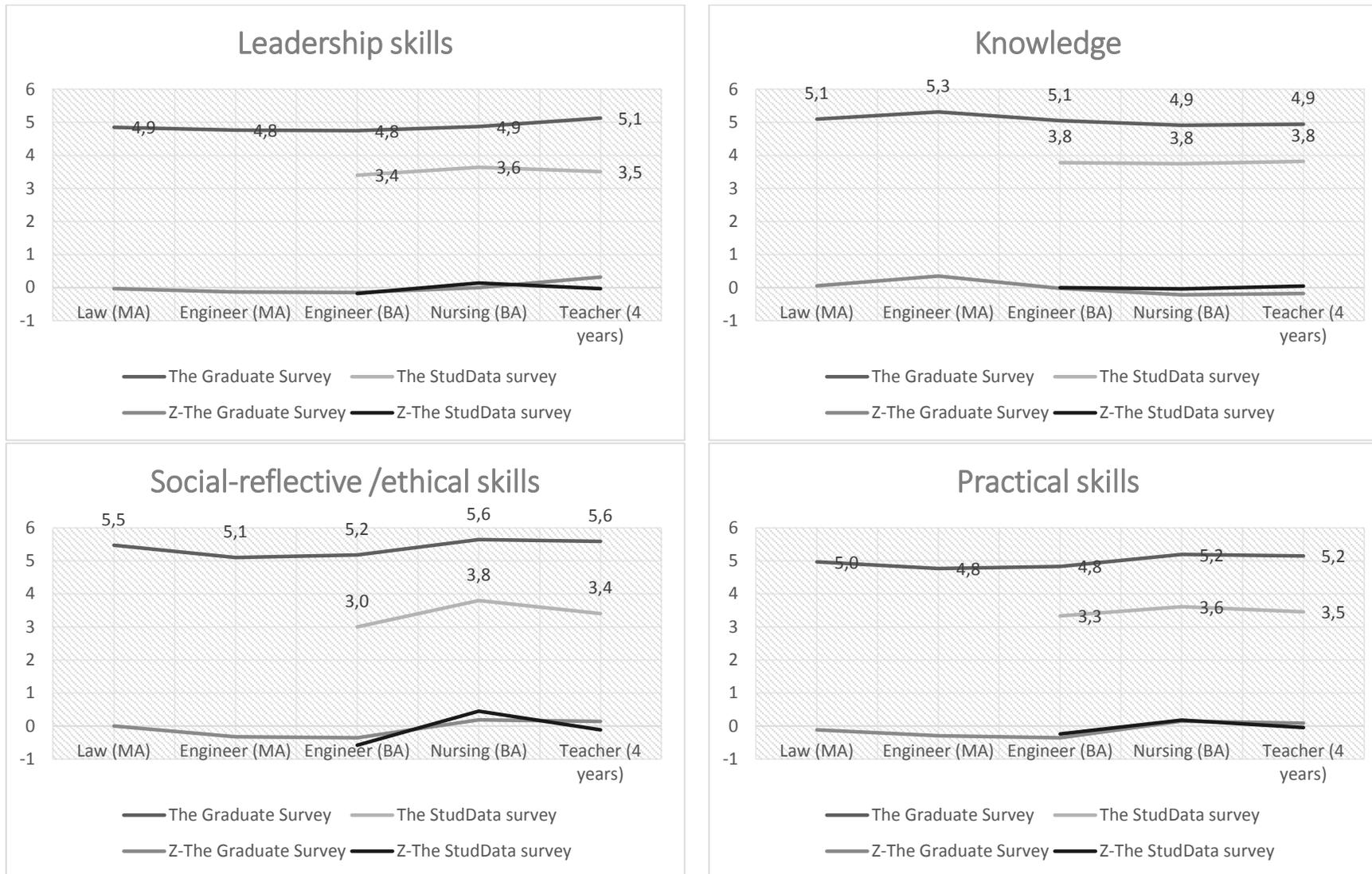


Figure 1. Leadership, social-reflective/ethical, practical skills and knowledge in the Graduate Survey and StudData. Mean scores and standardized mean scores (Z, mean=0, std. dev=1).

ⁱ Corresponding author: email: joakim.caspersen@nifu.no, telephone: +47 22 59 21 39. Fax: +47 22 59 51 01. Post: Joakim Caspersen, NIFU Nordic Institute for Studies in Innovation, Research and Education, PB. 5183 Majorstuen, 0302 Oslo, Norway.