

2019

Perceptions of Assessment in Norwegian Occupational Therapy Students

Tore Bonsaksen

OsloMet - Oslo Metropolitan University


Astrid Gramstad

UiT The Arctic University of Norway

Gry Mørk

VID Specialized University

Susanne G. Johnson

*Western Norway University of Applied Sciences*Follow this and additional works at: <https://encompass.eku.edu/jote> Part of the [Adult and Continuing Education Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Health and Physical Education Commons](#), and the [Occupational Therapy Commons](#)

Recommended Citation

Bonsaksen, T., Gramstad, A., Mørk, G., & Johnson, S. G. (2019). Perceptions of Assessment in Norwegian Occupational Therapy Students. *Journal of Occupational Therapy Education*, 3 (3). Retrieved from <https://encompass.eku.edu/jote/vol3/iss3/2>

This Original Research is brought to you for free and open access by the Journals at Encompass. It has been accepted for inclusion in Journal of Occupational Therapy Education by an authorized editor of Encompass. For more information, please contact Linda.Sizemore@eku.edu.

Perceptions of Assessment in Norwegian Occupational Therapy Students

Abstract

In addition to securing minimum standards of learning among students, assessment is increasingly used as a tool to improve students' learning. Assessment quality is measured as part of the Course Experience Questionnaire; however, the original 'appropriate assessment' scale has demonstrated ambiguous psychometric properties. The current study aimed to gain knowledge about occupational therapy students' perceptions of assessment, and this was achieved by examining the factor structure and internal consistency of the six items on the 'appropriate assessment' scale. Students from six Norwegian universities ($n = 187$, response rate 61.3 %) completed the scale and reported demographic information. The factor structure of the scale was assessed with Principal Components Analysis and Parallel Analysis, and internal consistency was assessed with Cronbach's α and mean inter-item correlations. Three factors were extracted from the data: Factor 1 (three items, $\alpha = 0.51$) is concerned with the content of assessment. Factor 2 (two items, $\alpha = 0.62$) is concerned with the context of assessment. Finally, Factor 3 (one item) is concerned with the feedback to students following assessment. In view of the item mean scores, feedback in particular appears to be an area of improvement for the occupational therapy programs, as well as a venue for further research.

Keywords

Assessment, feedback, higher education, learning environment, students

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Acknowledgements

The authors thank the participants who volunteered to take part in this study. In addition, they gratefully acknowledge the efforts made by Kjersti V. Helgøy (VID Specialized University), Vår Mathisen and Lene A. Åsli (UiT- The Arctic University of Norway), Linda Stigen, Trine A. Magne, and Tove Carstensen (Norwegian University of Science and Technology), who contributed to the data collection for this study.

Perceptions of Assessment in Norwegian Occupational Therapy Students

Tore Bonsaksen, MSc, OT ^{1,2}, Astrid Gramstad, PhD, OT ³, Gry Mørk, OT ², and

Susanne G. Johnson, MSc, OT ⁴

Oslo Metropolitan University, Norway¹

VID Specialized University, Sandnes, Norway²

University of Tromsø- The Arctic University of Norway ³

Western Norway University of Applied Sciences, Norway⁴

ABSTRACT

In addition to securing minimum standards of learning among students, assessment is increasingly used as a tool to improve students' learning. Assessment quality is measured as part of the Course Experience Questionnaire; however, the original 'appropriate assessment' scale has demonstrated ambiguous psychometric properties. The current study aimed to gain knowledge about occupational therapy students' perceptions of assessment, and this was achieved by examining the factor structure and internal consistency of the six items on the 'appropriate assessment' scale. Students from six Norwegian universities ($n = 187$, response rate 61.3 %) completed the scale and reported demographic information. The factor structure of the scale was assessed with Principal Components Analysis and Parallel Analysis, and internal consistency was assessed with Cronbach's α and mean inter-item correlations. Three factors were extracted from the data: Factor 1 (three items, $\alpha = 0.51$) is concerned with the content of assessment. Factor 2 (two items, $\alpha = 0.62$) is concerned with the context of assessment. Finally, Factor 3 (one item) is concerned with the feedback to students following assessment. In view of the item mean scores, feedback in particular appears to be an area of improvement for the occupational therapy programs, as well as a venue for further research.

INTRODUCTION

Assessment is an integral part of the higher education experience, and more than any other factor it influences what and how students learn (Boud, 1988; Tennant, McMullen, & Kaczynski, 2003). Students' qualifications are assessed before they are admitted to the study program, and they are assessed more or less regularly, for

different purposes, as they proceed with their study. Assessment serves two general purposes (Ramsden, 2003; Tennant et al., 2003; Wiliam, 2011). Assessment of learning, also called summative assessment, often involves grading and aims at assessing the student's level of knowledge or skill at the conclusion of a course. In outcomes-based education guided by the constructive alignment framework (Biggs & Tang, 2007), this form of assessment is directed towards the student's learning in relation to pre-established learning outcomes. On the other hand, assessment for learning, also called formative assessment, denotes assessment aimed at adding to and improving the student's learning (Wiliam, 2011). Thus, formative assessment is process oriented, whereas summative assessment is product oriented.

In recent years, formative assessment has attracted much interest and is perceived an important tool for learning. In Norway, this was highlighted in a recent white paper constituting the foundation for current debates on higher education (Department of Knowledge, 2016-2017). In this white paper, assessment is described as essential for students' learning and their continued motivation for studying. At the same time, this aspect of the students' higher education experience is one with which they continuously report to be less satisfied, compared to other aspects (Damen, Keller, Hamberg, & Bakken, 2016). For example, in retrospect, one in five higher education students in Norway report to be dissatisfied with the feedback he or she has received during the study program (Damen et al., 2016). This is in contrast to the high level of general satisfaction reported by higher education students in the country, and providing more frequent and higher quality feedback to students should therefore be focused within the higher education institutions (Department of Knowledge, 2016-2017).

Assessment has frequently been considered an aspect of the larger learning environment, or the 'presage' factors (Biggs, 1989). It has been suggested that the presage factors may affect learning outcomes directly or indirectly, in the latter case by influencing the students' approaches to learning (Biggs, 1989; Diseth, 2007; Lizzio, Wilson, & Simons, 2002). Ramsden (1991) found five learning environment factors to be derived from his proposed questionnaire, namely clear goals and standards, emphasis on independence, good teaching, appropriate workload, and appropriate assessment. By using scales to measure these concepts, a substantial amount of research has been in support of associations between learning environment factors and study behaviors and outcomes in a variety of student groups (Baeten, Kyndt, Struyven, & Dochy, 2010; Diseth, 2007; Diseth, Pallesen, Hovland, & Larsen, 2006; Gustin, Abbiati, Bonvin, Gerbase, & Baroffio, 2018; Kreber, 2003; Sharma, 1997; Wilson, Lizzio, & Ramsden, 1997), including occupational therapy (Richardson, Gamborg, & Hammerberg, 2005; Sadlo & Richardson, 2003). For example, Lizzio, Wilson, and Simons (2002) found that the higher-order measure of good teaching, which included the 'appropriate assessment' scale, was associated with a deep approach to studying and with greater satisfaction, greater generic skill development, and better grades. In Norway, Diseth and co-workers (2010) used structural equation modelling to demonstrate that approaches to studying mediated the relationships between perceptions of the learning environment and exam performance.

In the context of an ongoing longitudinal study concerned with learning processes and the learning environment among Norwegian occupational therapy students, we

found that the items included in the 'appropriate assessment' scale as developed by Ramsden (1991) did not yield an acceptable level of internal consistency (Carstensen et al.; Mørk et al.). The finding indicated that the scale items do not measure aspects of the same concept. Previous studies have indeed produced ambiguous results concerned with the properties of this particular scale, as internal consistency measures have ranged considerably (α between 0.47 and 0.74) between studies (Ramsden, 1991; Richardson, 1994; Wilson et al., 1997). In addition, two of the items concerned with staff learning from students (item # 6) and providing qualitative feedback in addition to grades (item # 29), have been shown to load more strongly on the 'good teaching' scale of the instrument. Thus, modified versions of the scale have been used in subsequent studies (Byrne & Flood, 2003; Ginns, Prosser, & Barrie, 2007). In view of the evidence suggesting assessment to contribute substantially towards learning (e.g., Baeten et al., 2010; Biggs & Tang, 2007; Boud & Falchikov, 2007; Richardson et al., 2005), and the emphasis currently placed on improving assessment in higher education (Damen et al., 2016; Department of Knowledge, 2016-2017), more exploration into the apparently multidimensional nature of assessment is needed.

Study Aim

The aim of the current study was to gain knowledge about occupational therapy students' perceptions of exams and assessments, as they had experienced them during the first half-year of their study program. This was accomplished by examining the structure of the responses they had provided to the 'appropriate assessment' scale.

METHODS

Design and Study Context

The study is a sub-study related to a longitudinal study of occupational therapy students in Norway, focusing on their perceptions of the learning environment and their approaches to studying. The present study had a cross-sectional design, employing data from the students while in their first year of study. The data were collected at a time of convenience at each of the universities, between December 2017 and March 2018.

Participants and Response Rate

Occupational therapy students at six higher education institutions in Norway were recruited for inclusion. From the six education programs, 305 students were eligible participants, and of these 187 students (response rate 61.3 %) participated. For each of the institutions, the response rates were $24/76 = 31.6$ % in Oslo, $56/77 = 72.7$ % in Trondheim, $19/39 = 48.7$ % in Gjøvik, $31/47 = 66.0$ % in Sandnes, $24/24 = 100.0$ % in Tromsø, and $33/42 = 78.6$ % in Bergen.

Measurement

Sociodemographic variables. Age (in years) and time spent on independent studying (average hours during a typical week) were registered as continuous variables. Gender, having prior experience from higher education, and having occupational therapy as the highest prioritized line of education at the time of enrolment were registered as categorical variables.

Appropriate assessment. The original Course Experience Questionnaire (CEQ; Ramsden, 1991) consists of 30 items distributed onto five scales: clear goals and standards, emphasis on independence, good teaching, appropriate workload, and appropriate assessment; the latter scale being used in this study. The concept 'appropriate assessment' refers to assessment and exam forms that promote and support students' learning processes. They emphasize understanding and independent reasoning, as opposed to the memorizing of facts. Appropriate assessments are distributed across the study course and aligned with the study progression, and are not only placed at the end of the course. They also emphasize formative feedback that goes beyond grading and brief, written feedback on assignments (Pettersen, 2007). One example item from the scale is: "Staff seem more interested in testing what you've memorized than what you've understood"; whereas all scale items are listed in Table 2. Modified and shorter versions of the 'appropriate assessment' scale have been employed in previous research (e.g., Byrne & Flood, 2003; Ginns et al., 2007); however, this study employed the already translated version of the original scale (Pettersen, 2007). Both in this translated version and the original, the items concerned with the 'appropriate assessment' scale are arranged throughout the CEQ instrument. All items were scored 1-5, reflecting a response continuum from disagreeing to agreeing with the item content. Scores were reversed prior to analysis, such that higher scores indicated little endorsement of each item's content.

Data Analysis

The sample was described with descriptive statistics; i.e., means and standard deviations for continuous variables and frequencies and percentages for categorical variables. Comparisons between men and women were performed using Chi-square tests (categorical variables) and independent *t*-tests (continuous variables).

When assessing latent factors in the 'appropriate assessment' scale items, an exploratory Principal Components Analysis (PCA) was performed. Bartlett's Test of Sphericity (Bartlett, 1954) was used to assess whether the variables' correlations were significantly different from zero. Expecting the scale items to be intrinsically correlated, we used the Direct Oblimin rotation method. Factor extraction was determined by inspecting the scree-plots, by assessing the Eigenvalue (λ) estimates, and by assessing the proportion of data variance explained by the factors. According to guidelines, factors with $\lambda > 1$ and/or factors explaining more than 10 % the variables' variance proportions were retained. In addition, we employed Parallel Analysis (Horn, 1965) in order not to overestimate the number of extracted factors (Zwick & Velicer, 1986). The Parallel Analysis proposes that factors should be retained only if the actual λ exceeds the randomly generated λ of the corresponding factor in a random dataset, when the same number of variables and respondents is used.

Statistical measures reported from the PCA include communalities, indicating the variance proportion of each variable explained by the factors together, and factor loadings, which are estimates of the association between a given variable and the extracted factors. To obtain a clearer view of the pattern, the factor loadings from the structure matrix were inspected, and loadings > 0.40 were considered high. Internal consistency was examined with Cronbach's α and inter-item correlations. Cronbach's

α coefficients exceeding 0.70 (Ponterotto & Ruckdeschel, 2007), and/or mean inter-item correlations exceeding 0.20 (Briggs & Cheek, 1986), were considered satisfactory.

Research Ethics

Approval for collecting, storing and utilizing the data was granted on October 12, 2017 by the Norwegian Center for Research Data (project no. 55875).

RESULTS

Participants

The characteristics of the study participants are displayed in Table 1. On average, the male participants were two years older than the female, and compared to the female participants, a larger proportion of the male participants had some experience from higher education from before they were enrolled in the occupational therapy program. Other than this, there were no systematic differences between men and women in the sample on the employed variables.

Table 1

Sociodemographic and Educational Characteristics of the Study Participants

Variables	All <i>n</i> = 187	Men <i>n</i> = 37	Women <i>n</i> = 149	<i>p</i>	<i>d</i>
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)		
Age	22.9 (4.6)	24.5 (5.5)	22.5 (4.3)	<0.05	0.41
Time spent on self-study	9.3 (7.0)	9.6 (7.9)	9.3 (6.8)	0.78	0.04
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)		
Had OT as priority education	117 (62.6)	24 (64.9)	93 (62.4)	0.78	
Had prior higher education	78 (41.7)	22 (59.5)	56 (37.6)	<0.05	

Note: One participant did not report gender. Time spent on self-study is average number hours during a typical week spent on independent study.

Item Scores

The participants' scores on each of the six 'appropriate assessment' scale items are displayed in Table 2. Five of the six items had a mean score between 3.5 and 3.8, whereas item #29 ("Feedback on student work is usually provided only in the form of marks and grades") received a considerably lower mean score ($M = 2.4$) in the sample.

Table 2

The Participants' Scores on the Six-item 'Appropriate Assessment' Scale

Items	n	M (SD)	Min-Max
6. Lecturers here frequently give the impression they have nothing to learn from students	183	3.7 (1.2)	1-5
10. To do well on this course all you really need is a good memory	185	3.8 (1.1)	1-5
15. Staff seem more interested in testing what you've memorized than what you've understood	186	3.8 (1.0)	1-5
26. Too many staff ask us questions just about facts	185	3.6 (0.9)	1-5
29. Feedback on student work is usually provided only in the form of marks and grades	185	2.4 (1.2)	1-5
32. It would be possible to get through this course just by working hard around exam times	184	3.5 (1.3)	1-5

Note: All item scores are reversed, i.e., higher scores indicate little endorsement of the statements' content. Items scores range between 1 and 5.

Factor structure and internal consistency of the scale. As the first step in the PCA, Bartlett's test of sphericity was found to be statistically significant ($p < 0.001$), indicating that the correlations between the scale items were significantly from zero. The items' communalities ranged between 0.51 (item 15) and 0.88 (item 29). Three factors had Eigenvalues > 1 and they were extracted: Factor 1, $\lambda = 1.82$, explaining 30.4 % of the data variance; Factor 2: $\lambda = 1.21$, explaining 20.2 % of the data variance; and Factor 3: $\lambda = 1.06$, explaining 17.7 % of the data variance. Thus, a cumulative 68.3 % of the total data variance was explained by the three factors together. When controlling the factor extraction with the Parallel Analysis, we found a randomly generated $\lambda = 1.03$ for Factor 3, which was lower than the actual λ found for the third extracted factor in the PCA. Thus, the Parallel Analysis confirmed that three factors should be extracted.

The structure matrix showed that all items loaded substantially (i.e., > 0.40) on one of the three factors, with no cross-loadings (i.e., all other factor loadings < 0.40). Three items loaded on Factor 1: Items #10 (need only a good memory), #32 (working hard around exam), and #15 (what you've memorized), factor loadings ranging between 0.61 and 0.76. Two items loaded on Factor 2: Items #6 (nothing to learn) and #26 (just about facts), with corresponding factor loadings 0.85 and 0.83. Finally, item #29 (only marks and grades) loaded on Factor 3 with a factor loading of 0.94. Factors 1 and 2 were positively correlated (0.18), whereas the correlations between Factors 1 and 3 (-0.07) and between Factor 2 and 3 were negative (-0.08).

The internal consistency of the three items belonging to factor 1 (items #10, #32, and #15) were Cronbach's $\alpha = 0.51$, mean inter-item correlation = 0.26. Deleting any of the items from this factor would lead to lower internal consistency between items. For the items belonging to Factor 2 (items #6 and #26), internal consistency was Cronbach's $\alpha = 0.62$, mean inter-item correlation = 0.46. As only one item (item #29) loaded on Factor 3, measures of internal consistency are irrelevant for this factor. The results from the PCA and the internal consistency analyses are displayed in Table 3.

Table 3

Factor Solution and Internal Consistency of the 'Appropriate Assessment' Scale (n = 179)

Item #	Factor 1	Factor 2	Factor 3	Comm.
10. Need only a good memory	0.76	0.11	-0.15	0.59
32. Working hard around exam	0.76	0.06	0.20	0.65
15. What you've memorized	0.61	0.36	-0.33	0.51
6. Nothing to learn	0.12	0.85	-0.18	0.74
26. Just about facts	0.16	0.83	0.14	0.74
29. Only marks and grades	-0.05	0.00	0.94	0.88
λ	1.82	1.21	1.06	
Cronbach's α	0.51	0.62	-	
Mean inter-item correlation	0.26	0.46	-	
Explained variance	30.4 %	20.2 %	17.7 %	
Total explained variance	68.3 %			

Note: Results derived from the exploratory Principal Component Analysis with Oblimin rotation and Kaiser normalization. Factor loadings are taken from the structure matrix. Comm. = communalities.

DISCUSSION

In this study we found that the six items comprising the 'appropriate assessment' scale, as derived from the Course Experience Questionnaire (Ramsden, 1991), empirically loaded on three factors. This appears to logically explain the low measure of internal consistency between the six items, a problem encountered in previous studies (Richardson, 1994) including studies from our own current project (Carstensen et al.; Mørk et al.). In essence, the findings suggest that assessment, as measured with the original 'appropriate assessment' scale, may be viewed as comprising three distinct aspects.

The three items belonging to Factor 1 (items # 10, # 32, and # 15; see Table 2 for item content) concentrated on what teachers in fact tend to assess, and what the student needs to do in order to perform well at exams. Thus, this factor is concerned with assessment content. The items are all phrased negatively and high endorsement of the items on this factor would indicate endorsing a view that assessment most of all requires memorizing and rote learning around exam time. Judging from the mean scores (Table 2), the sample as a whole trended towards disagreeing with the items' content, indicating a perception that assessment required a broader approach. However, all response options were used for each of the items and the correlations between the three items, as indicated by their uniform factor loadings (Table 3), indicate that the students had a stable pattern of responses to them: Those who tended to agree with one item on the factor tended also to agree with the other two. Conversely, those who tended to oppose the meaning content of one item on Factor 1 tended also to oppose the content of the other two.

It appears that Factor 1 is concerned with assessment of lower versus higher levels of learning, a conceptualization that readily aligns with Biggs' Structure of the Observed Learning Outcome (SOLO) taxonomy (Biggs & Tang, 2007). Students

agreeing with the items effectively state that the assessments, as they have experienced them, function largely as memory checking. Teaching and assessment aligned to this level of knowledge – oriented towards memorizing facts, rather than understanding relationships between facts – is likely to foster surface approaches to learning and poorer learning outcomes among students (Baeten et al., 2010; Lizzio et al., 2002; Richardson, 2005; Wilson et al., 1997), and should therefore be discouraged. Moreover, Tennant and co-workers (2003) reminded that assessment should not only concern the students' level of academic understanding, but also their know-how and skills in practical settings. The latter 'authentic assessment' type may be preferred within the healthcare institutions and services, and it would certainly involve assessing students' learning at the relational and generalizing levels of the SOLO taxonomy (Biggs & Tang, 2007). The three items included in Factor 1 may serve as an indicator of the students' perceived level of learning required in the relevant course.

The two items associated with Factor 2 (items # 6 and # 26; see Table 2 for item content) focus on the teachers' attitudes and preoccupations, as expressed by their behaviors toward the students. Thus, these items comprise aspects of the context for assessment. In comparison to the items loading on Factor 1, they are not concerned with the content of the assessment, and thus less directly concerned with what the exam situation requires of the student. However, the degree to which students perceive their teachers to have concern for more than plain facts would logically influence how they prepare for assessments.

Similarly, the degree to which students perceive teachers to have an interest in their reasoning, indicating that teachers might have something to learn from the students, is likely to impact on the students' sense of confidence and feelings about assessment in general. Again, judging from the mean scores of the items belonging to Factor 2 (Table 2), the sample as a whole trended not to agree with the items' content. However, all response options were used for the two items, indicating that there were students who tended to agree with a negative view of the assessment context.

It has been noted that the power differential related to assessment is more outspoken than in any other area of education (Nicol & Macfarlane-Dick, 2006). Thus, teachers' and supervisors' words – not to mention their absence of words, and their gestures and bodily expressions – may create an atmosphere that inhibits learning (Snyder, 2018). Creating curiosity and a positive climate for learning may be promoted by keeping focused on reasoning rather than facts, and by conveying a positive regard for students. The two items included in Factor 2 may therefore serve as an indicator of the learning context as perceived by the students.

Factor 3 was comprised by only one item (item # 29; see Table 2 for item content), which was concerned with feedback to the students. Importantly, as shown by the item mean score (Table 2), this was the one aspect of assessment towards which the sample held a more negative perception. This indicates that the students in fact were in relative agreement with the statement that feedback only occurred in the form of marks and grades given by the teachers. This form of assessment aligns with the notion of summative assessment (Boud & Falchikov, 2007; Scriven, 1991), a type of assessment needed to ascertain that students achieve the defined minimum

standards and to distinguish between students achieving better and poorer learning outcomes. Traditionally, most of the assessment resources in higher education have been directed towards the summative assessment type (Havnes & Raaheim, 2006). However, in view of research and policy emphasizing the role of formative feedback in student learning (Department of Knowledge, 2016-2017; Wiliam, 2011), the over-reliance of summative assessment, as indicated by the participants in this study, appears to be particularly problematic for the involved education programs. Rather, to add to the students' learning process, both the nature and the timing of assessment should be considered. Assessment directed towards enhancing the students' learning is quite different from assessment in the form of grades – generally, the first form serves the student, while the latter form serves the education program's systemic requirements. Formative feedback that is provided regularly throughout the education program can inspire, redirect, and challenge the student to move forward, while feedback in the form of final course grades merely states the quality of what the student has achieved (Wiliam, 2011). Students may also benefit from taking part in the assessment process themselves, by assessing their own work or performance, or that of their peers (Boud, 2000).

Interestingly, previous factor-analytic research on the assessment scale has found the item concerned with feedback to load more strongly on the factor named 'good teaching' (Ramsden, 1991). Consequently, we may envision that in the student's mind, receiving good feedback can feel like learning rather than undergoing assessment. In particular, when feedback is not just a response to what the student has produced, but also includes a concrete discussion on how the students may improve his or her performance, feedback – or rather: feed-forward (Boud & Falchikov, 2007), becomes teaching. For students, feedback is also vital on an existential level; without it, occupational therapy students undergoing fieldwork recently reported feeling "lost, scattered, scared, cautious, disappointed and angry [...]" (Snyder, 2018). Based on this particular study, it appears the item relating to feedback can somewhat crudely indicate the quality of feedback provided in the relevant course.

Considering how the 'appropriate assessment' scale items loaded on three factors, as discussed above, we suggest that assessment is a multi-faceted phenomenon. Moreover, we suggest that the facets as detected in this study largely relate to the content of assessment; the context for assessment; and the feedback following assessment. The students develop an understanding of the content knowledge they have to possess, reproduce, apply or synthesize to perform well at exams, and how they need to study to be able to do so. They develop this understanding within the context of teachers' general attitudes towards students and the curriculum. Finally, the students' experiences after exams and assessments will inform their views on what kinds of feedback they can expect following assessment. The interrelationships between these facets of assessment are modelled in Figure 1.

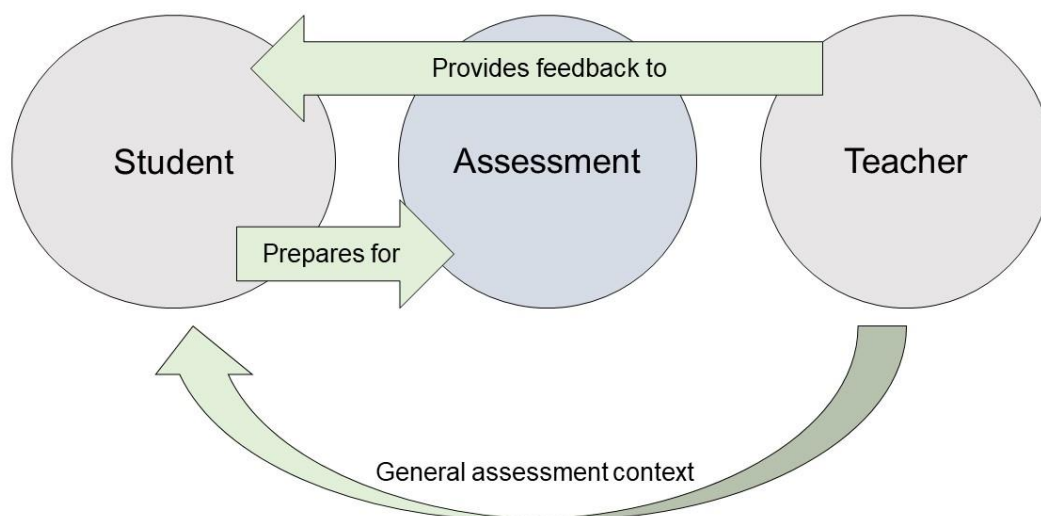


Figure 1. A model of three facets of assessment. The arrows represent the factors derived from the ‘appropriate assessment’ scale. Factor 1 is the students’ preparations for assessment and their perceptions of what is required to do well (assessment content). Factor 2 is the teachers’ general attitudes towards students and the curriculum (assessment context). Factor 3 is the teachers’ feedback on the students’ work (feedback).

Study Limitations and Future Research

The study is based on the use of self-report questionnaires only. Thus, the results show the structure of the students’ responses to the questionnaire items, yielding a view of the latent factors in the instrument. However, the thoughts and explanations of the students themselves, as pertaining to assessment, remain unknown. The scale consistency measures (Cronbach’s α) were lower than usually considered acceptable. However, a low number of items on scales, as was the case in this study, is commonly associated with low Cronbach’s α estimates (Ponterotto & Ruckdeschel, 2007). In view of the short scales, the acceptable mean inter-item correlations provide evidence of their internal consistency (Briggs & Cheek, 1986).

The sample was recruited from all of the six current occupational therapy education programs in Norway, and the sample size was sufficient for the employed analytic procedures. However, the response rate varied substantially between the universities, which may detract from our ability to generalize the findings to the larger population of undergraduate occupational therapy students.

Future studies may employ qualitative interviews with students so they can convey their views of assessment in their own words, in addition to responding to predetermined questions with fixed response options in questionnaires. Studies may be conducted longitudinally, or with different year cohorts, to examine whether views on assessment are stable over time, and whether responses are different or similar for less experienced compared to more experienced students.

CONCLUSION

The aim of the current study was to gain knowledge about occupational therapy students' perceptions of exams and assessments. We examined the structure of their responses to an 'appropriate assessment' scale, and found three latent factors: they concerned assessment content, assessment context, and feedback following assessment. In view of the low mean score on the item related to feedback, this appears to be an area of improvement for the occupational therapy programs, as well as a venue for further research.

References

- Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centered learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, 5(3), 243-260. <https://doi.org/10.1016/j.edurev.2010.06.001>
- Bartlett, M. S. (1954). A note on multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society*, 16(2), 296-298. <https://doi.org/10.1111/j.2517-6161.1954.tb00174.x>
- Biggs, J. (1989). Approaches to the enhancement of tertiary learning. *Higher Education Research and Development*, 8(1), 7-25.
- Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university*. Berkshire: UK: Oxford University Press.
- Boud, D. (1988). *Developing student autonomy in learning*. London: Kogan Page.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>
- Boud, D., & Falchikov, N. (Eds.). (2007). *Rethinking assessment in higher education. Learning for the longer term*. Oxon, UK: Routledge.
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54(1), 106-148. <https://doi.org/10.1111/j.1467-6494.1986.tb00391.x>
- Byrne, M., & Flood, B. (2003). Assessing the Teaching Quality of Accounting Programmes: An evaluation of the Course Experience Questionnaire. *Assessment & Evaluation in Higher Education*, 28(2), 135-145. <https://doi.org/10.1080/02602930301668>
- Carstensen, T., Stigen, L., Gramstad, A., Åsli, L. A., Mørk, G., Johnson, S. G., . . . Bonsaksen, T. Associations between study approaches, the learning environment, and occupational therapy students' academic performance (manuscript in progress).
- Damen, M. L., Keller, L. D., Hamberg, S., & Bakken, P. (2016). *Studiebarometeret 2015: hovedtendenser [The Study Barometer: Main Tendencies]*. Report no. 1. Retrieved from https://www.nokut.no/globalassets/studiebarometeret/2016/studiebarometeret_2015_hovedtendenser_1-2016.pdf
- Department of Knowledge. (2016-2017). *Meld. St. 16: Kultur for kvalitet i høyere utdanning [Culture for Quality in Higher Education]*. Oslo: The Government.
- Diseth, Å. (2007). Approaches to learning, course experience and examination grade among undergraduate psychology students: Testing of mediator effects and construct validity. *Studies in Higher Education*, 32(3), 373-388. <https://doi.org/10.1080/03075070701346949>

- Diseth, Å., Pallesen, S., Brunborg, G. S., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Higher Education*, 59(3), 335-352. <https://doi.org/10.1007/s10734-009-9251-8>
- Diseth, Å., Pallesen, S., Hovland, A., & Larsen, S. (2006). Course experience, approaches to learning and academic achievement. *Education and Training*, 48(2/3), 156-169. <https://doi.org/10.1108/00400910610651782>
- Ginns, P., Prosser, M., & Barrie, S. (2007). Students' perceptions of teaching quality in higher education: The perspective of currently enrolled students. *Studies in Higher Education*, 32(5), 603-615. <https://doi.org/10.1080/03075070701573773>
- Gustin, M. P., Abbiati, M., Bonvin, R., Gerbase, M. W., & Baroffio, A. (2018). Integrated problem-based learning versus lectures: a path analysis modelling of the relationships between educational context and learning approaches. *Medical Education Online*, 23(1), 1489690. <https://doi.org/10.1080/10872981.2018.1489690>
- Havnes, A., & Raaheim, A. (2006). Assessments and exams. In H. Strømsø, K.H. Lycke, & P. Lauvås (eds.), *When learning is most important. Teaching in higher education* [in Norwegian: Når læring er det viktigste. Undervisning i høyere utdanning] (pp.73-90). Oslo: Cappelen Akademisk Forlag.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. <https://doi.org/10.1007/BF02289447>
- Kreber, C. (2003). The relationship between students' course perception and their approaches to studying in undergraduate science courses: A Canadian experience. *Higher Education Research & Development*, 22(1), 57-75. <https://doi.org/10.1080/0729436032000058623>
- Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education*, 27(1), 27-52. <https://doi.org/10.1080/03075070120099359>
- Mørk, G., Magne, T. A., Carstensen, T., Stigen, L., Åsli, L. A., Gramstad, A., . . . Bonsaksen, T. Associations between the learning environment and students' approaches to studying: A cross-sectional study (manuscript in progress).
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>
- Pettersen, R. C. (2007). *Students' experience with and evaluation of teaching and the learning environment: Presentation of the Course Experience Questionnaire (CEQ) and validation of three Norwegian versions [in Norwegian: Studenters opplevelse og evaluering av undervisning og læringsmiljø: Presentasjon av Course Experience Questionnaire (CEQ) og validering av tre norske versjoner, Erfaringer med studiet (EMS)]*. Halden, Norway: Østfold University College. Report no. 4. <https://doi.org/10.7577/nova/rapporter/2005/19>
- Ponterotto, J. G., & Ruckdeschel, D. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills*, 105(3 Part I), 997-1014. <https://doi.org/10.2466/pms.105.3.997-1014>

- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education*, 16(2), 129-150. <https://doi.org/10.1080/03075079112331382944>
- Ramsden, P. (2003). Assessing for understanding. In P. Ramsden (Ed.), *Learning to teach in higher education* (2 ed., pp. 176-206). New York, NY: RoutledgeFalmer.
- Richardson, J. T. E. (1994). A British evaluation of the Course Experience Questionnaire. *Studies in Higher Education*, 19(1), 59-68. <https://doi.org/10.1080/03075079412331382143>
- Richardson, J. T. E. (2005). Students' perceptions of academic quality and approaches to studying in distance education. *British Educational Research Journal*, 31(1), 7-27. <https://doi.org/10.1080/0141192052000310001>
- Richardson, J. T. E., Gamborg, G., & Hammerberg, G. (2005). Perceived academic quality and approaches to studying at Danish schools of occupational therapy. *Scandinavian Journal of Occupational Therapy*, 12(3), 110-117. <https://doi.org/10.1080/11038120510030898>
- Sadlo, G., & Richardson, J. T. E. (2003). Approaches to studying and perceptions of the academic environment in students following problem-based and subject-based curricula. *Higher Education Research & Development*, 22(3), 253-274. <https://doi.org/10.1080/758482623>
- Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. Maclaughlin & D. C. Phillips (Eds.), *Evaluation and education: At a quarter century* (pp. 19-64). Chicago, IL: University of Chicago Press.
- Sharma, D. S. (1997). Accounting students' learning conceptions, approaches to learning, and the influence of the learning-teaching context on approaches to learning. *Accounting Education*, 6(2), 125-146. <https://doi.org/10.1080/096392897331532>
- Snyder, K. (2018). Exploring students' use of feedback during occupational therapy Level II fieldwork experiences. *Journal of Occupational Therapy Education*, 2(2). <https://doi.org/10.26681/jote.2018.020204>
- Tennant, M., McMullen, C., & Kaczynski, D. (2003). Assessment. In M. Tennant, C. McMullen, & D. Kaczynski (Eds.), *Teaching, learning and research in higher education: A critical approach* (pp. 97-110). New York, NY: Routledge. <https://doi.org/10.4324/9780203875919>
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the Course Experience Questionnaire. *Studies in Higher Education*, 22(1), 33-53. <https://doi.org/10.1080/03075079712331381121>
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432-442. <https://doi.org/10.1037/0033-2909.99.3.432>