Students’ perception of frequent assessments and its relation to motivation and grades in a statistics course: a pilot study

Bram E. Vaessen, Antoine van den Beemt, Gerard van de Watering, Ludo W. van Meeuwen, Lex Lemmens & Perry den Brok

To cite this article: Bram E. Vaessen, Antoine van den Beemt, Gerard van de Watering, Ludo W. van Meeuwen, Lex Lemmens & Perry den Brok (2017) Students’ perception of frequent assessments and its relation to motivation and grades in a statistics course: a pilot study, Assessment & Evaluation in Higher Education, 42:6, 872-886, DOI: 10.1080/02602938.2016.1204532

To link to this article: http://dx.doi.org/10.1080/02602938.2016.1204532

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 03 Jul 2016.

Submit your article to this journal

Article views: 731

View related articles

View Crossmark data

Citing articles: 1 View citing articles
Students’ perception of frequent assessments and its relation to motivation and grades in a statistics course: a pilot study

Bram E. Vaessen, Antoine van den Beemt, Gerard van de Watering, Ludo W. van Meeuwen, Lex Lemmens and Perry den Brok

Eindhoven School of Education (ESoE), Eindhoven University of Technology, Eindhoven, the Netherlands

ABSTRACT
This pilot study measures university students’ perceptions of graded frequent assessments in an obligatory statistics course using a novel questionnaire. Relations between perceptions of frequent assessments, intrinsic motivation and grades were also investigated. A factor analysis of the questionnaire revealed four factors, which were labelled value, formative function, positive effects and negative effects. The results showed that most students valued graded frequent assessments as a study motivator. A modest number of students experienced positive or negative effects from assessments and grades received. Less than half of the students used the results of frequent assessments in their learning process. The perception of negative effects (lower self-confidence and more stress) negatively mediated the relation between grades and intrinsic motivation. It is argued that communication with students regarding the purpose and benefits of frequent assessments could mitigate these negative effects.

Introduction

Assessment is considered one of the most important aspects of higher education, because students’ learning processes and learning outcomes are strongly influenced by the assessment system (Gibbs and Simpson 2004). Furthermore, educational institutions rely on assessment to qualify students for important certifications. Assessment is therefore pivotal for the existence of educational institutions (Meyer et al. 2010).

If well designed, assessment and resulting feedback can contribute to and improve the learning process (Black and Wiliam 1998; Hattie and Timperley 2007). If designed poorly, assessment can hinder learning by causing anxiety (Huxham, Campbell, and Westwood 2012), low self-esteem (Betts et al. 2009), low motivation for learning (Zilberberg et al. 2009) and inappropriate study behaviours such as surface approaches to learning (Scouller 1998) and cramming (Reeves 2006).

A commonly used method to counter the aforementioned problems is to use frequent (or continuous) assessment (Holmes 2015). Instead of a single final examination, multiple smaller assessment tasks are deployed, each counting towards the final grade. A meta-analysis by Bangert-Drowns, Kulik, and Kulik (1991) showed that frequent assessments can motivate students to study more regularly, resulting in reduced cramming and increased final grades. Frequent assessments can also lower anxiety (Fulkerson and Martin 1981), increase self-esteem (Cole and Spence 2012) and promote deep learning.
(McDowell et al. 2011; Carrillo-de-la-Peña and Pérez 2012). There are, however, several aspects that have to be taken into account to fully understand the impact of frequent assessment interventions.

First, the effects of any assessment system on students' learning processes and learning outcomes are mediated by their perceptions of this assessment system (Struyven, Dochy, and Janssens 2005; Lizzio and Wilson 2013). These perceptions are influenced by students' previous experiences with assessment, their characteristics and their communication with teachers and other students. This means that any intervention involving assessment can be perceived in various ways by students, and thus has differential effects on students' learning processes. Therefore, in order to fully understand the effects of such an assessment intervention, students' perceptions have to be taken into account.

Second, the exact details of the implementation of the frequent assessments are important. When frequent assessment is used, students will gain experience with the assessment system early in the course, altering their perceptions of the assessment within the course. Students use assessment as a direct source of the hidden curriculum (Sambell and McDowell 1998). In other words, students use their experience with assessment tasks to discover or alter their perceptions of what is expected of them. Inappropriate forms of assessment, for example measuring factual knowledge when understanding is the learning goal, can push students towards surface approaches to learning early in the course (Thomas and Bain 1984; Entwistle and Entwistle 1992). On the other hand, high quality assessment with appropriate feedback can encourage students to get involved in the material and use a deep approach to learning (Gibbs and Simpson 2004). When frequent assessment is deployed, the influence of the assessment system on students' study approaches increases, whether it be positive or negative.

Finally, assessment can influence students' psychological states, most importantly their motivation and self-efficacy. For frequent assessments this becomes more apparent, because students experience assessment early in the course. These experiences and resulting changes in psychological state may influence how students handle subsequent assessment tasks. Both positive and negative effects of assessment on motivation and self-efficacy may be self-reinforcing. For example, a student may have problems preparing appropriately for an assessment task due to low self-efficacy (van Dinther, Dochy, and Segers 2011). When this results in a low grade, the self-efficacy of the student diminishes even further, affecting preparation for the next assessment task. In a similar manner, good grades may increase motivation and self-efficacy, having positive effects on the preparation of the subsequent assessment task. Therefore, in research including frequent assessment, grades for frequent assessments can be studied as dependent variable as well as independent variable.

**Students’ perceptions of frequent assessment**

Recently, attention to students' perceptions of assessment has increased, but no consensus has been reached on what exactly constitutes these perceptions. The level at which the perceptions are measured differs between studies. Some research focuses on students' perceptions of specific assessment methods (Flores et al. 2015), or specific assessment tasks (Gulikers, Bastiaens, and Kirschner 2004). Others focus on students' perceptions of assessment in general (Brown and Wang 2013), or their perception of teachers' assessment behaviour (Alquraan 2014). The aspects of the measured perceptions also differ across research. Some examples are the perceived purpose of assessment (Hirschfeld and Brown 2009), the perceived fairness of assessment tasks (Flores et al. 2015), the perceived learning function of assessment in general (Brown and Wang 2013) or the enjoyment of assessment activities (Healy, McCutcheon, and Doran 2014), to name a few.

In the current study, we are interested in the perception of students regarding the specific intervention of frequent assessment. This implies that the level of measured perceptions is that of an ‘assessment system’, i.e. the combination of several planned assessment tasks within a single course. To our knowledge, no research has yet focused on the perceptions of this specific intervention. Therefore, a new instrument was developed and tested for this study. In developing the instrument items, we took into account the underlying rationales behind the intervention (Lemmens 2015), as well as the effects assessment can have on students.
The first rationale behind the intervention is that students can be taught the benefits of appropriate study behaviours (i.e. studying regularly, keeping up with the material), by extrinsically motivating them with graded frequent assessments that count towards the final grade. The relative weight of these frequent assessments is supposed to get lower in higher study years, because students are expected to have learned the benefits of studying regularly by then, and no longer need the extrinsic motivation to do so. The second rationale of the intervention is that each frequent assessment can be used formatively to help students in their learning process, by providing adequate feedback (Black and Wiliam 1998; Lemmens 2015).

Grades and intrinsic motivation

A much cited concern with graded assessments is the impact on intrinsic motivation (Butler and Nisan 1986; Pulfrey, Darnon, and Butera 2013), which is defined as engaging in a task because the activity itself feels rewarding. Self-determination theory (Deci and Ryan 2011) distinguishes intrinsic motivation from extrinsic motivation, which is defined as engaging in a task for external reasons. It is often found that intrinsic motivation for academic tasks is a good predictor of academic success (Linnenbrink and Pintrich 2002), and may therefore be considered an important outcome.

Intrinsic motivation for an assessment task is mainly influenced by how much that task contributes towards a person's feeling of autonomy and competence (Deci and Ryan 2011). Intrinsic motivation for a task is believed to be low when a person feels controlled to engage in the task, or the task does not provide a feeling of competence. A problem with giving grades for assessment tasks is that students perceive grades as an external reward or punishment. Students may feel controlled by that reward/punishment in their decision to engage with the task. In other words, students may perceive that the reason they engage in the assessment task is because they have to obtain a passing grade, or because they strive for the reward of a high grade, instead of engaging in the task out of interest. Because the decision to engage feels controlled, the assessment task does not provide the need for autonomy, lowering students' intrinsic motivation for the task.

Most research on rewards and intrinsic motivation (Butler and Nisan 1986; Deci, Koestner, and Ryan 1999; Pulfrey, Darnon, and Butera 2013) focuses on isolated tasks in a laboratory setting, with the explicit assumption that the 'interest value of the target tasks was at least neutral' (Deci, Koestner, and Ryan 1999, 635). In higher education, however, assessment tasks are not isolated. They are part of a course, and it could be argued that intrinsic motivation for a course as a whole is more important than intrinsic motivation for assessment tasks. Whether the feeling of autonomy for assessment tasks influences intrinsic motivation for the course as a whole is not known. Also, the initial interest value of assessment tasks in higher education may be low. In that case, the effect of providing grades for assessment tasks may not influence intrinsic motivation much.

An aspect that has not received much attention yet is the relation between the height of grades and intrinsic motivation. With graded frequent assessment, students receive grades frequently during the course, and the height of these grades may affect students' intrinsic motivation. Moreover, the relation between the grades received and intrinsic motivation may be mediated by students' perceptions of frequent assessment. The current study investigates students' perceptions of autonomy regarding frequent assessment, as well as the relations between grades, perceptions and intrinsic motivation.

The research question that guides our investigation is: What aspects of students' perceptions of frequent assessments can be distinguished, and how are these perceptions related to grades for frequent assessments, intrinsic motivation, and final grades?

Method

Context

This research was performed in a statistics course for the Industrial Engineering bachelor programme in a technical university in the Netherlands. In this bachelor programme, students learn to examine business processes from the perspectives of people, engineering and organisation. Mathematics and
statistics play an important role in this programme, for example to model, analyse and optimise business processes.

The statistics course introduces basic concepts of statistics such as probability distributions, estimation theory, regression models and confidence intervals. It is somewhat tailored to technical engineering, for example by emphasising certain probability distributions that are common in business modelling, or using appropriate contexts during classes, assignments and assessment. It is an obligatory course, scheduled in the last quarter of the first year of the bachelor programme. The course lasted about nine weeks, and, in a typical week, 6 h of plenary lectures and 2 h of ‘instruction classes’ were planned. In the instruction classes, students could work on practice items and ask questions.

The assessment in this course consisted of three parts: weekly assessments, an assignment on regression analysis and the final examination. The weekly assessments were given in the last half hour of the instruction classes, and were taken online on students' laptops. These assessments covered the topic of the previous week, and consisted of five short-answer parameterised items that usually involved calculations, and required filling in the correct results. For some items, a relevant industrial engineering context was given. Students immediately received the results of the assessment after they completed it. In the evening, the students could consult the correct answers with short explanations online. In total, six of these weekly tests were given. The first was a practice test, and from the other five, the best four grades counted. Students who did not do the weekly test received a zero.

In the week before the final examination, students did a 1.5 h open-book assignment-based assessment on the topic regression analysis. The assignment involved analysing data and interpreting the results of these analyses, for which they had to write and hand in a report. After the assignment, a model solution was made available online, and, as part of the assessment, students needed to write a short reflection on their own work, comparing it to the model solution, and identifying possible mistakes they made. The report, together with the reflection, was used to grade the assignment, which could be a 0, 2.5, 5, 7.5 or 10.

A 3-h cumulative final examination was given in the eighth week, which consisted of open ended questions with industrial engineering contexts. The weekly assessments together with the assignment on regression analysis counted for 30% of the final grade, while the final assessment counted for the remaining 70%.

Participants
From the 319 students who enrolled in the class, 219 students participated in this research (66% response rate). The participants’ age ranged from 18 to 34 (mean = 20, SD = 1.54; 73% male). Group differences between students who participated in the study and those who did not show that participants were significantly younger with a small effect size (d = .24), and scored statistically significantly higher on their final grade with a small to medium effect size (d = .40), when compared to non-participants. Therefore, there seem to be a selection bias for students who performed well in the course, which could be partially mediated by engagement in the course. No gender differences were found between participants and non-participants.

Measurements
Students’ perceptions of frequent assessment
The items for measuring students’ perceptions of frequent assessments were designed by the researchers, based on the intentions of the intervention, and on several research findings, which are listed in Table 1. The items were checked and revised by two educational experts, which resulted in some minor language revisions. Each item was answered with a five-point Likert scale ranging from totally disagree (1) to totally agree (5). A principal component analysis (PCA) with varimax rotation resulted in five factors, of which four were interpretable. The items from the uninterpretable factor (items 5 and 18) were removed from analysis.
Intrinsic motivation

To measure intrinsic motivation for studying in general, the scale ‘taking initiative’ from the Dutch PROSDL questionnaire (Van Meeuwen et al. 2012) was used. The items in this scale were ‘I regularly do extra work for my study, purely out of interest’, ‘During my study I take the initiative for learning new things, instead of waiting for my teacher to encourage me’ and ‘Even when a course has ended, I continue to learn about the subject’ (translations from Dutch by authors). These items measure ‘free-choice time on task’, which is an often used measure for intrinsic motivation (Deci, Koestner, and Ryan 1999).

To measure intrinsic motivation for studying for the course, three items were used from the intrinsic motivation scale of the academic self-regulation questionnaire (Vansteenkiste et al. 2009). The items are ‘I studied with pleasure for this course’, ‘Studying for this course was boring’ and ‘I found studying for this course to be interesting’ . These items measure intrinsic motivation by the self-reported enjoyment of studying for the course, another often used measure for intrinsic motivation (Deci, Koestner, and Ryan 1999).

A PCA with varimax rotation confirmed that the six questions measure the two separate constructs as intended, with all six items having high factors loadings (≥.80) on one factor and low factor loadings.
(≤.17) on the other factor. The reliability is good for both intrinsic motivation for studying in general ($\alpha = .77$) and intrinsic motivation for the course ($\alpha = .87$).

**Other measures**

The mean of the grades of the five weekly assessments and the practical assessment on regression analysis was used as a measure of the grades for frequent assessments. Furthermore, the grade for the final examination was used, and students’ Grade Point Averages (GPAs) were used as covariate.

**Procedure**

The questionnaire for this study was embedded in the course evaluation form. Students retrieved an invitation to fill out this form after the final examination was taken and before they received their grade on the final examination.

**Analyses**

Factor analysis on the 20 items of the perceptions of frequent assessments questionnaire was performed by means of a PCA with varimax rotation. The factor scores were calculated as the mean of the items related to each factor. Cronbach’s Alpha or Spearman-Brown Coefficient (for the two-item scale) was used for reliability analysis of the factors. Two-tailed Pearson’s product-moment coefficients were used to test for correlations. To test the mediating role of perceptions of frequent assessments in the relation between grades and intrinsic motivation, path analysis via regression was used. All variables were entered directly in the model. To test whether perceptions of frequent assessments could predict examination grades, a hierarchical regression analysis was used, where the first model only consisted of the covariates GPA and the motivation variables, and the second model also included the perception variables.

**Results**

**Students’ perception of frequent assessments**

In Table 2, the factor loadings of a four factor solution for the 18 items on students’ perception of frequent assessments are shown, ordered by factor, then by factor loading.

**Factor 1 (value)**

This factor can be interpreted as the ‘value’ students attach to the frequent assessments. Some students may value frequent assessments because they believe frequent assessments are an impulse to study regularly and stay motivated, while other students attribute low value to frequent assessments because they think assessments are a waste of time and that they would be better off without them. The reliability of this factor is good ($\alpha = .86$). Note that the factor loadings of the items in Table 2 would result in a reversed measurement of value.

**Factor 2 (negative effects)**

Two of the items in this factor refer to negative consequences that arise from frequent assessments, namely a drop in confidence and stress. The third item asks students whether they would prefer frequent assessments without grades. The relation between these three items could be explained by test anxiety and low grades. Students may experience anxiety and stress for subsequent assessments when they obtain a low grade for an assessment. These students may therefore prefer ungraded assessments. The reliability of this factor is acceptable ($\alpha = .71$).
Factor 3 (formative function)

We interpret the items in this factor to reflect the ‘formative’ function of assessment (Black and Wiliam 1998). Students can use grades and feedback from frequent assessments as a source of information to adjust their learning processes. Higher scores on this factor means students’ learning processes were more influenced by frequent assessments. The reliability of this factor is questionable ($\alpha = .65$).

Factor 4 (positive effects)

The items in this factor are interpreted as being indicators for positive effects of frequent assessments, namely more self-confidence and less tension, somewhat the opposite of the ‘negative effects’ factor. Item 14 had a low item-total correlation, and its content was less related to the other two items, therefore this item was removed from the factor. The reliability of this factor is poor (Spearman-Brown Coefficient = .54). Despite the poor reliability, the factor seemed theoretically interesting as a direct contrast to the ‘Negative Effects’ factor, therefore it was included in the analyses. Results of analyses that included this factor should be interpreted with caution.

Descriptive statistics

Students’ perception of frequent assessment

Table 3 shows the descriptive statistics for the scores on the four factors of the students’ perceptions of frequent assessments. The value of the frequent assessments was evaluated as high ($M = 3.59$), with 73% of the students responding positively, and only 16% responding negatively. Overall, then, students...
valued the frequent assessments positively. The scores on formative function were average. About 44% of the students reported that they used frequent assessments in their learning process. In general, scores on positive ($M = 2.61$) and negative ($M = 2.52$) effects were low. About 15% of the students reported that they experienced stress and loss of self-confidence from frequent assessments, whereas 10% of the students responded that they experienced reduced stress and increased confidence because of the frequent assessments.

Table 4 shows the correlations between the factor scores for the different perceptions of frequent assessments. The value students attach to frequent assessments showed a moderately negative correlation with negative consequences, and a moderately positive correlation with formative function and positive effects. Furthermore, a small to moderate positive correlation was found between formative function and positive effects. Finally, there was a small to moderate negative correlation between positive effects and negative effects.

**Motivation**

Table 5 shows the descriptive statistics of intrinsic motivation for studying in general (IM general) and intrinsic motivation for the course (IM course). The mean scores for intrinsic motivation for studying in general (2.72) is less than the ‘neutral’ score 3, which shows that intrinsic motivation for studying in general is low. About 15% of the students reported they did extra (non-assessed) studying purely out of interest. Intrinsic motivation for the course scored slightly higher among the participants than the neutral 3. The correlation between these two measurements was moderate, $r = .25$.

**Grades**

Table 6 shows descriptive statistics of grades for frequent assessments (GradeFA) and final examination (GradeExam), as well as of students’ GPA. The GPA was calculated from the final grades of all previous courses finished. Grades for the final examination ($M = 4.18$) were notably lower than grades for frequent assessments ($M = 6.14$). The correlations between the measurements were high between GradeFA and GradeExam ($r = .64$), high between GPA and GradeExam ($r = .69$) and moderate between GPA and GradeFA ($r = .49$).

**Correlations**

Table 7 shows the correlations between the variables used in this study.

**Students’ perception of frequent assessments, grades for frequent assessments and motivation**

Table 8 shows the correlations between perceptions, grades and motivation. As expected, the grades students received for frequent assessments were moderately related to the perception of these frequent assessments. Higher grades were related to higher perception scores for value, lower perception scores for negative effects and higher perception scores for positive effects. This means that students with high grades perceived the effects of frequent assessments as more positive (lower stress and more confidence), and valued frequent assessments more. Students with lower grades perceived negative effects (more stress and less confidence) and valued frequent assessments less.
Some of the perceptions of frequent assessments were related to intrinsic motivation for the course. As expected, a moderate negative correlation was found between perceptions of negative effects and intrinsic motivation, and a small positive correlation was found between positive effects and intrinsic motivation. There was also a small correlation between grades and intrinsic motivation.

The results of the path analysis, which tested the mediating role of perceptions of frequent assessments in the relation between grades and intrinsic motivation, are shown in Figure 1. The top part of

Table 4. Correlations between perceptions of frequent assessments (n = 210).

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Formative function</td>
<td>.35***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Negative effects</td>
<td>−.35***</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>4. Positive effects</td>
<td>.42***</td>
<td>.25***</td>
<td>−.26***</td>
</tr>
</tbody>
</table>

***p < .001.

Table 5. Scores on intrinsic motivation variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M (SD)</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM general</td>
<td>216</td>
<td>2.72 (.77)</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>IM course</td>
<td>216</td>
<td>3.05 (.87)</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 6. Students’ grades.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M (SD)</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>209</td>
<td>6.30 (.77)</td>
<td>3.66</td>
<td>7.88</td>
</tr>
<tr>
<td>GradeExam</td>
<td>207</td>
<td>4.18 (2.55)</td>
<td>.00</td>
<td>9.90</td>
</tr>
<tr>
<td>GradeFA</td>
<td>210</td>
<td>6.14 (1.65)</td>
<td>.00</td>
<td>9.60</td>
</tr>
</tbody>
</table>

Table 7. Correlation matrix for all variables in this study.

<table>
<thead>
<tr>
<th>Value</th>
<th>Negative effects</th>
<th>Formative function</th>
<th>Positive effects</th>
<th>IM general</th>
<th>IM course</th>
<th>GPA</th>
<th>GradeFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative effects</td>
<td>−.35**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative function</td>
<td>.35**</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive effects</td>
<td>.42**</td>
<td>−.26**</td>
<td>.25**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM general</td>
<td>−.04</td>
<td>−.07</td>
<td>.13</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM course</td>
<td>.08</td>
<td>−.26**</td>
<td>.05</td>
<td>.19**</td>
<td>.27**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>.07</td>
<td>−.24**</td>
<td>−.04</td>
<td>.14*</td>
<td>.15*</td>
<td>.17*</td>
<td></td>
</tr>
<tr>
<td>GradeFA</td>
<td>.25**</td>
<td>−.35**</td>
<td>−.10</td>
<td>.37**</td>
<td>.04</td>
<td>.18**</td>
<td>.49**</td>
</tr>
<tr>
<td>GradeExam</td>
<td>.20**</td>
<td>−.30**</td>
<td>−.05</td>
<td>.29**</td>
<td>.03</td>
<td>.24**</td>
<td>.69**</td>
</tr>
</tbody>
</table>

Note: n = 207.
*p < .05
**p < .01.

Table 8. Correlations between students’ perception of frequent assessments, grades for frequent assessments and intrinsic motivation for the course.

<table>
<thead>
<tr>
<th>Value</th>
<th>Negative effects</th>
<th>Formative function</th>
<th>Positive effects</th>
<th>GradeFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GradeFA</td>
<td>.25**</td>
<td>−.35**</td>
<td>−.10</td>
<td>.37**</td>
</tr>
<tr>
<td>IM course</td>
<td>.07</td>
<td>−.26**</td>
<td>.04</td>
<td>.19**</td>
</tr>
</tbody>
</table>

Note: N = 210.
**p < .01.
Figure 1. Path analysis for perceptions of frequent assessments as mediator between grades on frequent assessments and intrinsic motivation for the course. Top: direct effect of grades on intrinsic motivation. Bottom: effect of grades on intrinsic motivation, mediated by perceptions of frequent assessments. The numbers represent the beta-coefficients, the * denotes significant predictors.

Table 9. Results of the regression analysis on intrinsic motivation for the course.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>t</th>
<th>p</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>−.13</td>
<td>.09</td>
<td>−.11</td>
<td>−1.37</td>
<td>.17</td>
<td>[−.32, .06]</td>
</tr>
<tr>
<td>Positive effects</td>
<td>.11</td>
<td>.08</td>
<td>.11</td>
<td>1.39</td>
<td>.17</td>
<td>[−.05, .26]</td>
</tr>
<tr>
<td>Negative effects</td>
<td>.24</td>
<td>.07</td>
<td>−.25</td>
<td>−3.29</td>
<td>.00</td>
<td>[−.38, −.10]</td>
</tr>
<tr>
<td>Formative function</td>
<td>.11</td>
<td>.09</td>
<td>.10</td>
<td>1.27</td>
<td>.21</td>
<td>[−.06, .29]</td>
</tr>
<tr>
<td>Grades</td>
<td>.05</td>
<td>.04</td>
<td>.09</td>
<td>1.17</td>
<td>.24</td>
<td>[−.03, .12]</td>
</tr>
</tbody>
</table>

Note: $R^2 = .09$.

Table 10. Results of the hierarchical regression analysis on final grades.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>t</th>
<th>p</th>
<th>F</th>
<th>$R^2$</th>
<th>Δ$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75.70***</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>1.64</td>
<td>.17</td>
<td>.50</td>
<td>9.53</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM general</td>
<td>−.31</td>
<td>.16</td>
<td>−.09</td>
<td>−2.01</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM course</td>
<td>.39</td>
<td>.14</td>
<td>.13</td>
<td>2.73</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GradeFA</td>
<td>.61</td>
<td>.09</td>
<td>.37</td>
<td>7.10</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38.61***</td>
<td>.60</td>
<td>.01</td>
</tr>
<tr>
<td>GPA</td>
<td>1.68</td>
<td>.18</td>
<td>.50</td>
<td>9.60</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM general</td>
<td>−.30</td>
<td>.16</td>
<td>−.09</td>
<td>−1.90</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM course</td>
<td>.34</td>
<td>.15</td>
<td>.11</td>
<td>2.29</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GradeFA</td>
<td>.53</td>
<td>.10</td>
<td>.32</td>
<td>5.52</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>.15</td>
<td>.20</td>
<td>.04</td>
<td>.75</td>
<td>.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive effects</td>
<td>−.11</td>
<td>.16</td>
<td>−.04</td>
<td>−.68</td>
<td>.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative effects</td>
<td>−.01</td>
<td>.19</td>
<td>.00</td>
<td>.05</td>
<td>.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative function</td>
<td>.20</td>
<td>.16</td>
<td>.07</td>
<td>1.30</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***p < .01.

Figure 1 shows the direct effect of grades on intrinsic motivation, which is .18 ($p < .05$). The percentage of variance explained is 3%. The bottom part of Figure 1 shows that, when perceptions of frequent assessments are tested as mediator, there is no longer a significant direct effect between grades and intrinsic motivation; rather there is an indirect path between grades and motivation through negative effects, $F(5209) = 4.44, p < .01$. With the mediators, the explained variance increased to 10%. Details of the results of the regression analysis are shown in Table 9.

Students’ perception of frequent assessments and final grades

Table 10 shows the results of the hierarchical regression analysis testing whether perceptions of frequent assessments could predict examination grades. In the first step, all covariates were significant predictors,
and they explained 59% of the variance. In the second step, all covariates except intrinsic motivation for studying in general were significant, and no perceptions variables were significant predictors. The inclusion of the perception variables in the model did not lead to a significant increase in the explained variance of examination grades. This means that if motivation, GPA and grades for frequent assessments are already taken into account, the perceptions of frequent assessments do not contribute to predicting grades on the final examination.

Conclusions and discussion

The first aim of this study was to explore aspects of perceptions of frequent assessments. Factor analysis showed that four aspects of perceptions of frequent assessments could be distinguished. The first aspect, labelled value, refers to how much value students attribute to frequent assessments. The majority of the students in this study valued frequent assessments. It did not seem to bother students that these frequent assessments controlled them in their studying behaviour; instead, students seemed to welcome the extrinsic motivation to study regularly. Lemanski (2011) also found that students favour assessment as a motivator for studying.

The findings on students’ perceptions of value should be interpreted within their context. In the current research group, students showed little intrinsic motivation for studying in general. Furthermore, it is known that many students dislike or even fear statistics courses (Gal and Ginsburg 1994; Onwuegbuzie and Wilson 2003). Therefore, most students likely have little or no initial intrinsic motivation for a statistics course, and the detrimental effect of grades on intrinsic motivation (Butler and Nisan 1986; Pulfrey, Darnon, and Butera 2013) is expected to be minimal. In other (non-obligatory) courses that are more closely related to students’ interests, intrinsic motivation may be more important, and students may value graded frequent assessments less due to their controlling nature. This means that teachers should take into account the nature of a course and students’ initial intrinsic motivation when designing assessments. When the initial intrinsic motivation is high, it might be better to use assessments that do not feel controlling (i.e. non-obligatory and feedback only). However, it is then important to help students engage with these formative assessments. A large study across many disciplines shows that most students ‘did not value, complete or even notice the presence of formative assessment tasks’ [referring to non-graded optional assessment tasks] (Jessop, Hakim, and Gibbs 2014, 77).

Furthermore, the value students attribute to frequent assessments was related to the grade they received for these assessments. Higher grades for frequent assessments were linked to higher perceptions of the value of frequent assessments. It could be that students’ perceptions of frequent assessments become biased by grades, due to an emotional response to either high or low grades (Carless 2006). Another explanation could be that students who attribute high value to frequent assessments are more likely to study hard for these assessments and therefore receive higher grades. To control students’ emotional responses to grades, and to increase their effort in studying for assessment, teachers should clarify to students that frequent assessments are intended as a valuable learning tool, and that grades are a reflection of their effort and current understanding rather than a reflection of their personal ability (Vispoel and Austin 1995; Brown, Peterson, and Irving 2009).

The second aspect of students’ perceptions of frequent assessments found in this study was labelled formative function. This aspect reflects whether students believed they had used the results from frequent assessments in their learning processes. Literature on formative assessments shows that learning from assessments (the formative function of assessment) is the most important and powerful aspect of assessment (Black and Wiliam 1998). In our study, however, less than half of the students reported to use frequent assessments as an aid for their learning process, even though there was ample opportunity to do so. The grades that students received for the frequent assessments may have decreased the use and effectiveness of feedback (Butler and Nisan 1986; Lipnevich and Smith 2008; Shute 2008). Students may be too much focused on the grade and discard the opportunity to learn from feedback. Teachers could consider making reflection on assessment feedback an obligatory part of their classes (Nicol and Macfarlane-Dick 2006), to ensure that assessments serve a formative function for all students.
Another approach that can increase students’ engagement with feedback, called *adaptive release of grades*, involves giving feedback to students prior to providing them with a grade (Parkin et al. 2012).

The perception of formative function was found to not predict grades on the final examination. It was expected that students who use feedback from frequent assessments in their learning would benefit and that their final examination grade would reflect this (Black and Wiliam 1998). However, no evidence was found in this study. Several possible explanations can be discerned. One explanation could be a lack of alignment between frequent assessments and the final examination. In the current study, however, an overlap existed between frequent assessments and the final examination. Another explanation could be that the accuracy of students’ judgement of their formative use of frequent assessments is low. Finally, it could also be that most students restudy all materials prior to the examination, which could negate the effects of the formative use of frequent assessments.

The third and fourth aspects of the students’ perceptions were *positive effects* and *negative effects*. These aspects reflect the perceived impact of frequent assessments on self-confidence and stress: either less stress and more self-confidence for *positive effects*, or more stress and less self-confidence for *negative effects*. A modest number of students reported either negative or positive effects from frequent assessments. Furthermore, a path analysis showed that the perceived negative effects of frequent assessments mediated the relation between grades and intrinsic motivation. Thus, low grades for frequent assessments diminish intrinsic motivation only when students perceive the frequent assessments as stressful and detrimental to their self-confidence.

It is important to try to mitigate this possible negative effect of frequent assessments. If students have enough intrinsic motivation to engage in non-graded assessment tasks, it may be better to only provide feedback, without a grade. However, if grades are necessary as a motivator, students should be encouraged to perceive frequent assessments as a valuable learning opportunity, even if their grade was low.

There are several limitations to the current study. First, the context of this study was limited to one course with students from one major and a selective group of students. As argued earlier, the motivation for studying in general was low for the students in this major, and the course we studied was an obligatory statistics course. Students from other majors, or from another course, could have different motivations, and perceptions of graded frequent assessments would be different as well. Nevertheless, we believe this pilot study has revealed interesting relationships between perceptions of assessment, grades and motivation, which may be generalisable to other contexts.

Another limitation concerns the psychometric properties of the questionnaire used for measuring perceptions of frequent assessments. The questionnaire is an experimental first version, and one of the scales showed insufficient reliability. Furthermore, two items did not load on any factor. However, three of the four factors did seem to work well, and, overall, the factors found in this study appear valid; they seem to match theoretical expectations about perceptions on frequent assessments, and most factors showed the expected relations with other measured constructs. There may, however, be other relevant aspects of perceptions of frequent assessments that we did not include.

Finally, there was a limit to how many questionnaire items were allowed to be added to the course evaluation forms. Therefore, we chose to only measure intrinsic motivation, while it might have been interesting to include data on extrinsic motivation as well. Other constructs that might be interesting to include in future research are the learning approaches students used, the effort they put into studying and the distribution of effort over time.

For future research, it might first be interesting to compare these results with results from similar research in non-obligated courses, and from other majors. Also, the long-term effect of learning could be incorporated. Research shows that frequent assessment influences how students spread their study effort over time (De Paola and Scoppa 2011), and while the effect on final grades may be limited, there is evidence that studying regularly, instead of cramming may benefit long-term retention of study materials (McIntyre and Munson 2008).
To conclude, the current pilot study showed the relevance of including students' perception of assessments in research on assessment. Students' perceptions of assessments can provide valuable insight into how assessment influences students. In regard to the debate on grades and intrinsic motivation (Butler and Nisan 1986; Pulfrey, Daron, and Butera 2013), we argue that there are multiple solutions, and that the nature of the course and student characteristics should be taken into account. Communication with students about the purpose and benefits of frequent assessments could be the key to reducing possible negative effects from graded frequent assessments and increasing students' effort on non-graded frequent assessment.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


