



## Teacher and student perceptions of intermediate assessment in higher education

Indira N. Z. Day, F. M. van Blankenstein, P. Michiel Westenberg & W. F. Admiraal

To cite this article: Indira N. Z. Day, F. M. van Blankenstein, P. Michiel Westenberg & W. F. Admiraal (2018) Teacher and student perceptions of intermediate assessment in higher education, Educational Studies, 44:4, 449-467, DOI: [10.1080/03055698.2017.1382324](https://doi.org/10.1080/03055698.2017.1382324)

To link to this article: <https://doi.org/10.1080/03055698.2017.1382324>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 27 Sep 2017.



Submit your article to this journal [↗](#)



Article views: 1837





View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)

## Teacher and student perceptions of intermediate assessment in higher education

Indira N. Z. Day<sup>a</sup> , F. M. van Blankenstein<sup>a</sup>, P. Michiel Westenberg<sup>b</sup> and W. F. Admiraal<sup>a</sup> 

<sup>a</sup>ICLON, Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands;

<sup>b</sup>Department of Psychology, Leiden University, Leiden, The Netherlands

### ABSTRACT

Universities introduce intermediate assessment because it is understood to have positive effects on student behaviour and achievement. Yet, how intermediate assessment is perceived might be conditional for its success. The current study investigates both teachers' and students' perceptions of intermediate assessment. Teachers and students were interviewed and Student Evaluations of Teaching were examined. Results indicate that both teachers and students had generally positive perceptions of intermediate assessment. However, the two groups provided different reasons for their positive perceptions. Teachers and students agreed that intermediate assessment helps students to keep up with their study work. Moreover, teachers mentioned that they could assess various knowledge and skills with intermediate assessment, whereas students preferred intermediate assessments to test the same knowledge and skills as the final exam. This finding suggests that teachers and students in our study had different goals for intermediate assessment.

### ARTICLE HISTORY

Received 10 June 2016

Accepted 16 September 2017

### KEYWORDS

Intermediate assessment;  
teacher perception; student  
perception; higher education

Universities often take measures to entice students to start studying earlier in the semester, to make completing study programmes more feasible, and to increase study success in general. One of these measures is intermediate assessment (van Berkel, Jansen, and Bax 2012). Intermediate assessment, often also referred to as continuous or frequent assessment, refers to assessment or assignments that take place during the course period. Measures to increase study success, like intermediate assessment, are often only regarded as successful when they improve academic achievement or pass percentages. The perceptions of stakeholders, like teachers and students, are often overlooked in this process. It can be argued that successful implementation of intermediate assessment hinges on its perceived usefulness. Several authors (e.g. Entwistle 1991; Prosser, Walker, and Millar 1996) have indicated that student perceptions of the learning context play an important role in student learning. van Dinther et al. (2014) found that perceptions predict self-efficacy, which in turn predicts

**CONTACT** Indira N. Z. Day  [i.n.z.day@iclون.leidenuniv.nl](mailto:i.n.z.day@iclون.leidenuniv.nl)

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

student outcomes. Since intermediate assessment can take various forms, it can be hypothesised that teacher and student perceptions are influenced by the type of assessment that is used. This study investigates both teachers' and students' perceptions of different forms of intermediate assessment.

### **Intermediate assessment in higher education**

In the current study, intermediate assessment is defined as an assessment which takes place during the course period, which can take many different forms, and is handed in to the teacher. As was mentioned before, intermediate assessment can have many different forms, like quizzes, essays, presentations, or projects for example. Other characteristics that can differ between intermediate assessments are for example the frequency, amount of questions, time to finish, modality (written/oral/computer/etc.), grading and weight of the assessment. Furthermore the quality of the assessment is of importance as well. Lizzio, Wilson, and Simons (2002) reiterated the importance of perceived appropriateness of the assessment for student learning approaches and student results. Intermediate assessments can be used for summative purposes, for formative purposes, or both. Voluntary homework assignments and class readings do not count as intermediate assessment. Findings from research into the effects of assessment in general and intermediate assessment in particular indicate both positive and negative effects. These will be discussed in more detail below.

### **Positive effects of intermediate assessment**

Intermediate assessment can lead to increased time-on-task as evidenced in a study by Admiraal, Wubbels, and Pilot (1999), where law students in a problem method course with several assessments spent more time studying than their peers in the traditional case-based course. Doumen, Broeckmans, and Masui (2014) demonstrated a relationship between increased study time and higher exam grades, but Michaels and Miethe (1989) discovered that only students who spread their study time during the semester, as opposed to students who resort to cramming before the exam, profit from increased study time. Since students often focus their learning activities around examinations, intermediate assessments are an incentive for students to spread their study time (Cohen-Schotanus 1999).

Higher results for students who spend more time studying during the semester may be explained by the cognitive advantages of spreading. Laboratory studies have shown that spreading content while studying leads to better retention in long-term memory than cramming does (Kornell 2009). The advantages of spreading, or distributed practice, have also been extensively reviewed by Dunlosky et al. (2013) who conclude that it is one of two learning techniques they rate highest.

The second high-rated learning technique is practice testing (Dunlosky et al. 2013). Whereas practice testing is usually undertaken as a homework activity by students, intermediate assessment works following a similar mechanism of regular testing. The effectiveness is caused by the "testing effect", which indicates that information is retained in memory better by testing it (Roediger and Karpicke 2006). The testing effect is explained by the advantages of effortful retrieval, that is to say, by trying to recall knowledge from memory, the retention of that memory is strengthened. Roediger and Karpicke (2006) reviewed a large body of testing effect studies and found support for the testing effect in laboratory studies as well as in authentic learning environments.

Intermediate assessment creates opportunities for feedback, especially when teachers implement formative intermediate assessments (assessments with a main goal of improving learning by feedback instead of judging achievement, Bennett 2011). Gibbs and Simpson (2004) propose that feedback improves learning especially when it is provided “just in time” (i.e. at a moment when it is still useful for students). Providing feedback to an intermediate assessment can improve exam results, as indicated by Bouwmeester and colleagues (2013). They found that students who participated in voluntary online formative assessments scored higher on midterms and final exams, compared to their peers who did not participate.

### ***Negative effects of intermediate assessment***

Intermediate assessment can also have adverse effects. Harland et al. (2015) concluded that assessing students frequently for grades led to strategic behaviour, because students were mainly focused on getting grades and not on gaining a deeper understanding of the material. Furthermore, competition arose, with non-graded study activities losing out. Harland et al. (2015) called this an “assessment arms race”. The notion of an assessment arms race is backed up by findings from Vos (1992) who found that students in the Netherlands spent a maximum of seven hours on studying. When the workload exceeded these seven hours, competition between study activities would occur.

Harland et al.’s (2015) conclusions with regard to the lack of deep understanding of the material may especially be problematic when there is a large overlap in content between the intermediate and final assessment. Research by Kling et al. (2005) indicates that frequent testing was only beneficial with regard to final exam scores when there was overlap between test items in the intermediate and final assessments. However, using several intermediate assessments with the same items as in the final exam is more in line with teaching to the test (for a definition see Popham 2001), and probably does not lead to deep learning.

Within the context of universities, perceptions of intermediate assessment are not always positive. An example of this is a report from our university’s newspaper where the following point is raised: Using a lot of intermediate assessments in higher education can make students feel patronised, and treated like they are back in high school (Bongers 2015).

Finally, intermediate assessment requires teachers to invest time in designing more assessments, and to provide more feedback and grades, than with final exams only (Leeming 2002). Additionally, when intermediate assessments are administered during class time, this reduces the time available for teaching. However, Leeming (2002) argued that this reduced teaching time mainly led to less redundancy of materials and not to omission of teaching core materials.

### ***Perceptions of intermediate assessment***

The aforementioned literature reflects both positive and negative effects of intermediate assessment in higher education. However, when discussing the potential and actual effects of intermediate assessment, these studies were silent about perceptions. We believe that teacher perceptions of assessment are an important factor in the implementation process. Several researchers have investigated assessment perceptions of teachers and students; however, these studies often focus on assessment in general and not on intermediate assessment in particular.

Goos, Gannaway, and Hughes (2011) used surveys, focus groups and semi-structured interviews to investigate the assessment perceptions of associate deans, course coordinators and students, showing that these three groups identify different challenges when discussing assessment. Course coordinators' comments focused a lot on what can be described as assessment conditions: workload, bureaucracy and lack of support. Conversely, students' main point of dissatisfaction seemed to be insufficient feedback, even though first year students were unable to qualify what sufficient feedback meant to them. Associate deans also identified providing feedback as the most pressing assessment issue.

Myers and Myers (2015) also found indication that teaching workload influenced teachers' choices when designing assessments. Based on a national survey among postsecondary teachers, they found that teachers who taught more classes were less likely to use learner-centred assessment, a term Myers and Myers use to describe "the assessment component of effective pedagogy" (1905). However, Myers and Myers suggest that teachers see workload and learner-centred assessment as a trade-off, which may explain why some teachers with higher teaching loads do employ learning-centred assessments and others do not. This individual trade-off can be influenced by teachers' assessment perceptions.

## Current study

Previous studies investigated assessment perceptions with assessment as a general construct. The current study aims to get insights into the perceptions of *intermediate* assessment. The main focus of this paper is teacher perceptions of intermediate assessment as these perceptions are crucial for implementing assessments in the first place. Student perceptions are also investigated, as intermediate assessment is implemented with student goals in mind. The following two research questions were formulated:

- (1) What types of intermediate assessment are used in the programmes under investigation?
- (2) How are these types of intermediate assessment perceived by teachers and students?

## Method

### Context

The research was conducted at the Law School (Bachelor programmes of Law and Criminology) and the Faculty of Social Sciences (Bachelor programmes Psychology and the International Bachelor of Psychology, IBP) of a research university in the Netherlands. Every university programme in the Netherlands consists of 60 European Credits (EC; 1680 study hours) per year, in the current programmes divided into 5 and 10 EC courses (running for a half or a full semester). Psychology and Law programmes traditionally used to have a limited number of contact hours, mainly in the form of large-scale lectures, and a lot of time for independent study. Over the past few years, the focus has shifted towards more contact hours and intermediate assessments to keep students on track. These programmes usually offer three concurrent courses, with one weekly lecture and one weekly or biweekly tutorial.

The programmes have obligatory intermediate assessments in the first-year courses, but the preconditions are very different. At the Law School, in courses with a study load of 10 EC, teachers are obliged to offer a partial exam, which counts toward the students' final grade. In 5 EC courses, teachers are free to choose whether they want to offer intermediate assessments. For both Psychology programmes, all students are graded on two different types of output: a final multiple choice exam which makes up 70% of the course grade, and an intermediate assessment, usually taking place in the workgroup sessions, for the other 30%. Apart from these preconditions, set by the institutional boards, all teachers are free to find a type of intermediate assessment that fits their course. Students are obliged to complete the assessments that are offered in their courses.

### **Participants**

Twenty-two first-year courses (eleven at each department) taught in the 2013–2014 academic year were selected for the study. Teachers were invited for interviews. To ensure that the interviews would have information about the lectures as well as the workgroups, for some courses pairs of teachers consisting of a lecturer and a workgroup teacher were invited. In total, 28 teachers were invited for interviews.

Eighteen teachers (10 males) representing fifteen courses accepted the invitation, ten teachers did not reply to the invitation or declined participation, because of time constraints, or because they felt they could not offer valuable insights. The average teaching experience of the participating teachers was 5.3 years (SD 3.6 years) of teaching the specific course they were interviewed on.

The number of full-time first-year students enrolled in the programmes was 830 (Law, including Fiscal and Notarial Law), 145 (Criminology), 480 (Psychology) and 120 (IBP). Information about student participation is included in the Data section.

### **Data**

The university's electronic study guide (e-prospectus) provides an overview of all courses offered. Courses that used intermediate assessment were selected. For the selected courses, detailed information from the study guide was analysed to gain insight into course details and intermediate assessments.

### **Teachers' perceptions**

**Interviews.** Semi-structured interviews using a pre-determined topic list were conducted by the first author. Teachers were asked about their general opinion on intermediate assessment, the plans they had for intermediate assessment and their reasons for deciding on these plans. Furthermore, the interview focused on the teachers' experiences using intermediate assessment in their course and their evaluations of it. Teachers were also asked about student opinions on their course and about their ideas on the results from Student Evaluations of Teaching (SET, for more information see below).

Teacher interviews were conducted individually or in pairs, took place at their workplace and lasted between 30 min and one hour. All interviews were audiotaped and the interviewer took notes during the interview.

### *Students' perceptions*

***Student evaluations of teaching (SET).*** To gain insight into the student opinions on the courses, relevant questions from Student Evaluations of Teaching (SET) were selected. The SETs were familiar to students and used the same terminology that students were used to hear in class. For Psychology SETs, three questions were chosen: (1) students rating their opinion on the course on a scale of 1 to 10 (with 1 = very poor and 10 = excellent); (2) students rating the workload of the course given the number of ECs on a five-point scale (with 1 = much too light, 3 = in proportion and 5 = much too heavy); and (3) students rating on a five-point scale whether they agreed that the assignments, including the intermediate assessment, were instructive (with 1 = disagree and 5 = agree). The latter question was not obligatory and was therefore not answered in all SETs. For the Law School, two questions were selected: (1) students' general opinion of the course on a five-point scale (with 1 = inadequate and 5 = very good), and (2) an open ended question where students could give a more detailed account of their opinions.

All SETs were administered anonymously. SETs of the Psychology and Criminology programmes were administered as hard copy evaluations after the final exam, SETs of the Law programme were online evaluations. No courses were evaluated both online and in hard copy. The response rate on the online evaluations was much lower than on the hard copy equivalents. The response rates to SETs is included in Tables 2 and 3.

***Interviews.*** To get a more detailed account of the SET scores, a small sample of Psychology students was recruited for an interview using the department's online research participation system (SONA System) and flyers in the department building. One female Psychology student was interviewed in June 2014 and six students (1 male) from the IBP were interviewed during the 2014–2015 academic year.

Semi-structured student interviews focused on how students perceived intermediate assessment in their courses. Students first gave their general opinion on intermediate assessment and discussed general study behaviour. Subsequently, they reported what the intermediate assessment in a course was, what they felt about the assessment and what they felt about the workload. Whereas teachers were interviewed about the one course they taught, students provided information about all courses they had taken up until that point. Student interviews were conducted individually on campus and took about half an hour.

### *Ethics*

The current study was approved by the ethics committee of the Psychology Department. All students signed informed consent forms before they were interviewed.

### *Analysis*

#### *Interviews*

All interviews were transcribed and structured using Atlas.ti 7 by the first author. Themes were used as a structuring tool to group quotes in general themes. We decided against a predetermined coding scheme to be able to stay close to the actual data, because of our small sample size. A preliminary set of five teacher interviews was structured inductively to get a feeling for the information in the interviews.

Subsequently, the first, second and fourth author decided on an open scheme consisting of general factors. After the general scheme with factors and preliminary themes had been set, all interviews were structured. New themes were created inductively during this process but fitting within the original factors. When five interviews had been structured using the new scheme, all the themes were reviewed critically by the first author and themes that overlapped were merged or redefined. The preliminary set of structured interviews was restructured as well.

The scheme consisted of four factors. The first factor was Content, where themes have a direct relation to the content of the studies. Example themes from the Content factor were “course” (all information directly related to the course, such as design, planning, exams, etc.) and “student” (information related to student opinions, behaviour, results, etc.) The second factor was Nature. The Nature factor applied to the type of information being discussed, such as “opinion” or “attitude”. The third factor was Judgement, which denoted whether the information discussed was “positive” or “negative”. The final factor was Not Relevant, where all information not directly related to the interview was grouped. Examples for this factor were “other programmes” and “general conversation”.

After structuring the teacher interviews, the scheme consisted of 56 themes, with 37 themes in the Content factor, seven in the Nature factor, six in the Judgment factor and six in the Not Relevant factor as well. These factors and their specific themes were not mutually exclusive, that is, different factors could be combined in one quote, to indicate a students’ positive opinion, for example.

Student interviews were structured using the same scheme. To accommodate for the student experiences, three themes specific to the student situation were added to the Content factor. After structuring all interviews, therefore, the scheme consisted of a total of 59 themes, with 40 themes in the Content factor.

Quote length was determined by the content of the quote, as long as the same topic was being discussed by the participants and the interviewer it counted as one quote. As a result, quotes differed in length. All information in the interviews was structured.

To assess the replicability of the analyses the second author, who had been involved in designing the scheme but not in structuring the interviews, investigated two main themes from the results. First, he investigated one interview to see whether all problems with the assessment that were discussed by the teacher were in fact categorised as problems by the first author. Differences between the first and second author were discussed to further the quality of the scheme. Second, he explored all quotes that were discussing workload to see whether the results that were drawn from these quotes were an accurate representation of the opinions expressed by the teachers. He agreed with the conclusions drawn by the first author, and substantiated the conclusions with information from the quotes.

## SET

To investigate differences in SET scores between the different assessment types, paired t-tests were run. These paired t-tests were run separately for students in the regular Psychology bachelor and the International Bachelor of Psychology, since students from both programmes have different backgrounds. The Law and Criminology programmes overlap in some courses, therefore it was not possible to run separate paired t-tests for these programmes.



## Results

### *Use of intermediate assessment*

Table 1 indicates that there is a wide variety in types of intermediate assessment across the different courses. Furthermore, several teachers employed different types of assessment within their course. In the teacher interviews, two main reasons for this diversity in assessment methods were mentioned. First, teachers mentioned that students need to learn a diverse set of skills in their course, and choosing diverse assessments is connected with that. Second, teachers reported that they wanted to experiment with different methods. Almost all teachers mentioned the relationship between the content of their course and the type of assessment as one of their main rationales for deciding on a specific type. The two teachers of Law course 5 described their decision for assessment methods as follows:

Teacher 1: Since we had no experience with more activating methods, we also thought, well let's try a few different things to see what works.

Teacher 2: Yes, and of course it is connected with the subject matter you're working on that week (Law 5)<sup>1</sup>

In about half of the courses the intermediate assessment was new. Based on the assessment information as presented in the study guide, the courses were clustered into four types: (1) only written assignments as intermediate assessment; (2) a mix of different assessment types; (3) a partial exam, in combination with another form of intermediate assessment; (4) a voluntary trial exam.

### *Teacher perceptions*

The results can be grouped in three categories: (1) general opinion on intermediate assessment; (2) conceptual remarks to improve the intermediate assessment; and (3) practical difficulties with intermediate assessment. Teachers who taught courses with different assessment types did not display large differences in experience. For this reason, the results will be discussed for all types together.

### *General opinion on intermediate assessment*

All teachers reported positive perceptions of intermediate assessment. Several reasons were mentioned. Ten teachers reported that intermediate assessment encouraged students to actively engage with the course content during the course period, which they in turn hoped would lead to more active participation in class. For example:

[...]the intermediate assessment is very good for different reasons, namely to force or encourage them to really start working on the subject-matter because the deadline of the assessment is earlier than the exam[...]. (Law 9)

Furthermore, teachers in six courses mentioned the added value of being able to test at multiple occasions and to measure a variety of knowledge and skills.

One of the teachers in a course that had a partial exam was critical of the preconditions surrounding the partial exam in her course, but felt very positive about the written assignments that students handed in, which she felt led to more intelligent questions and participation in class. Another teacher mentioned that intermediate assessments can make a course more manageable for students and introduce them to the demands of the course.

**Table 1.** Overview of first year courses that use intermediate assessment.

Course	EC	Final exam	Intermediate assessment	Details of assessment	N	Feedback	Grade	Group
<b>Law 1</b>	10	MC questions	Partial exam, open ended <sup>a</sup>	Essay question based on a case study	1		20%	3
			Digital MC trial test	Course content	12	Automated		
			Workgroup assignment <sup>a</sup>	Case study	12	Personalised once		
<b>Law 2</b>	5	Essay questions	Experiment & Report <sup>a</sup>		1		Pass/fail	2
			Proposition <sup>a</sup>		2		Pass/fail	
<b>Law 3</b>	5	MC questions	Duo presentation <sup>a</sup>		1		Pass/fail	1
			Workgroup assignment <sup>a</sup>		1	Peer feedback		
<b>Law 4</b>	10	Essay question						
		Open questions	Portfolio <sup>a</sup>	Nine written workgroup assignments	1		30%	2
		MC questions		e.g. APA writing, observation, essay				
<b>Law 5</b>	5	MC questions	Presentations	Portfolio assignments	6			2
			Workgroup assignment <sup>a</sup>	Includes Case, Plea, Debate				
<b>Law 6</b>	10	Essay question(s)						
		Essay questions	Partial exam, open ended <sup>a</sup>	Essay questions on first half of the course	1		50%	3
<b>Law 7</b>	5	MC questions	Group presentation <sup>a</sup>		1		Pass/fail	2
			Workgroup assignment <sup>a</sup>	Includes case, Essay,	7			
<b>Law 8</b>	5	Essay question						
		MC questions	Workgroup assignment <sup>a</sup>	Essay	1	Peer feedback		2
		Essay question		Practice exam questions	4	Plenary feedback		
<b>Law 9</b>	5	Open questions	Paper <sup>a</sup>	Conduct and report on an interview in pairs	1	Peer feedback	30%	2
<b>Law 10</b>	5	Written exam						
			Workgroup assignment <sup>a</sup>		3	Plenary feedback		1
<b>Law 11</b>	5	Final Paper	Written assignments <sup>a</sup>					
		Self-reflection			5	Peer feedback		1
						Personalised once		

(Continued)

**Table 1.** (Continued).

Course	EC	Final exam	Intermediate assessment	Details of assessment	N	Feedback	Grade	Group
<b>Psychology 1</b>	5	MC questions	Short essay test <sup>a</sup>		8		30%	1
<b>Psychology 2</b>	5	MC questions	Digital MC trial test		1			4
<b>Psychology 3</b>	5	SPSS skills test	Digital MC trial test		1			
<b>Psychology 4</b>	5	MC questions	Workgroup assignment <sup>a</sup>	Short essay, debate	4		30%	2
			Digital MC trial test		1			
			Essay <sup>a</sup>		3	Peer feedback	30%	1
<b>Psychology 5</b>	5	MC questions	Digital MC trial test		1			
			Workgroup assignment <sup>a</sup>	Cases, MC questions	3		30%	2
				Poster	1	Peer feedback		
<b>Psychology 6</b>	5	MC questions	Digital MC trial test	Presentation in pairs,	1		30%	2
			Workgroup assignment <sup>a</sup>	paper	2			
<b>Psychology 7</b>	5	Two choice questions	Digital MC trial test		1			
			Behavioural change paper <sup>a</sup>					
<b>Psychology 8</b>	5	MC questions	Digital MC trial test				30%	2
			Digital MC trial test					4
<b>Psychology 9</b>	5	SPSS skills test	Digital MC trial test					
			Digital MC trial test					4
<b>Psychology 10</b>	5	MC questions	Practical assignments <sup>a</sup>	Analyse and interpret experiments	4		Pass/fail	2
				Write research report on one experiment	1		Pass/fail	
<b>Psychology 11</b>	5	MC questions	Test questions <sup>a</sup>		4		30%	1

Notes: Law denotes the Law School, the Law 1 and 7 courses are part of the Bachelor of Law. The Law 2, 4, 6, 9, 10 and 11 courses are part of the Bachelor of Criminology. The Law 3, 5 and 8 courses are part of both programmes. Psychology courses are the same for the IBP and the regular Bachelor Psychology. Courses in bold indicate teachers participated in an interview. N = frequency of an assessment.

<sup>a</sup>Mandatory intermediate assessment.

Their positive perceptions notwithstanding, teachers also had quite a few criticisms. A few of these considerations were conceptual, others were of a more practical kind.

### *Conceptual criticisms*

The majority of the teachers did not mention that the intermediate assessment could lead to exam training or loss of student responsibility. Only three teachers argued that they believed the goal of intermediate assessment should not be to endlessly practise and make the courses too easy for students. Conversely, one teacher mentioned actually wanting to practise the skills for students even more.

Some teachers were critical of the preconditions that were set for the intermediate assessment. One of the teachers of a course with an obligatory partial exam reported: “That it’s obligatory, yeah I don’t think that’s strictly necessary” (Law 6). Her opinion was shared by another teacher who was not obliged to use an intermediate assessment but was opposed to compulsory assessments. Other criticisms of the preconditions were related to the fact that students got two partial grades that they could use to compensate. The possibility for compensation may be especially problematic for courses where two very different sets of skills or knowledge were tested. One teacher said “I think that it actually should not be possible, because we want them to be able to do both and not one or the other by chance” (Psychology 2). This teacher and one of his co-workers also felt that the distribution of 70% final exam and 30% intermediate assessment was giving too much weight to the latter.

### *Practical considerations*

Teachers mentioned keeping their own workload in mind while deciding on an assessment type. Both the Psychology and the Law programme have a large student body, whereas the student population of the Criminology programme is moderate in size.

The workload encountered by the teachers was different for each course. Teachers in courses with personalised feedback and grades reported a heavier burden of workload than the teachers of other courses. As Table 1 shows, three courses did not have graded assignments or individualised feedback, so the only time investment for the teachers was to check if students had completed the assessment. However, in one of the Law courses one teacher was responsible for grading 150 portfolios, a massive workload. Two teachers mentioned the burden of the workload but also reported that it was worth it.

In one course with graded assignments, the teacher mainly reflected upon the workload involved in setting up the whole system of assessment and grading. However, he did comment that “[...] of course that’s done now, it was a one-time investment” (Psychology 1). Five courses used peer feedback to alleviate teacher workload.

In addition to the workload directly related to the assessment, teachers acknowledged that there was an increased workload in administration and communication. One example of this was dealing with students who missed an assessment, especially since missing too many assessments often led to exclusion from the final exam.

Five teachers discussed plagiarism and their solutions to prevent this. Because of the large student body, the work group meetings of the courses were dispersed over the week. This meant that students in a Monday work group could let their peers in a Friday work group know what the assessment questions would be. Teachers tried to overcome this problem in several ways. In three courses the deadline for completing the assessment was the same regardless of the timing of the work group, which sometimes led to criticism from students,

for example because the deadline was before the lecture discussing the topic of the assignment. Another teacher invested a lot of time in designing different questions for each work group. Yet another teacher made all assignment questions available beforehand but found that students posted the correct answers on social media, which he hoped to prevent next year by increasing the number of possible questions.

Another practical consideration that came up during three of the interviews was re-sits of the intermediate assessment. According to institution guidelines, students had the right to re-sit a failed exam, but in one course the intermediate assessment did not count separately. If a student failed the final exam they had to retake the intermediate and final assessment as one set, even if they had passed the intermediate assessment. Additionally, in two courses the re-sit had a different form from the intermediate exam, which teachers feared might lead to strategic thinking on the part of students:

But then they can decide [...] instead of going to a work group eight times, I'll just prepare the book, take two exams, one multiple choice, one essay exam, because that way I can pass as well. (Psychology 1)

When asked if they wanted to change things in the assessment, teachers mainly addressed fine-tuning the current assessment practice and not reshaping it entirely. Some teachers wanted to eliminate specific aspects they felt did not work as planned. One teacher specifically mentioned that it would be inadvisable to change the intermediate assessment right away: "Look, we need to consider this. You shouldn't immediately overturn things. So we need to continue this structure" (Psychology 6).

### ***Student perceptions***

#### ***Student evaluations of teaching***

Tables 2 (Psychology) and 3 (Law and Criminology) show the SET results of the two different departments. Results will be discussed per department.

Results from the paired samples t-tests comparing assessment types, indicate that students did not evaluate courses differently depending on the type of intermediate assessments administered (all  $p$ 's > 0.05). This was apparent in the regular Bachelor of Psychology as well as the IBP. Independent of the assessment types, students rated most courses as 6 (out of 10) or higher, indicating a moderately positive evaluation.

At the Law School, paired t-tests also showed no differences between the three types of assessment (all  $p$ 's > 0.05), again implying that courses were not evaluated differently solely based on the type of intermediate assessment that was used. As can be seen in Table 3, all courses except for one were rated 3.5 (out of 5) or higher, indicating a generally positive evaluation.

#### ***Students' detailed accounts on SET***

The Psychology students provided three explanations in the interviews which flesh out their generally positive evaluations of intermediate assessment. First, just like teachers, students valued the fact that the intermediate assessment kept them on track. Second, they believed that completing the intermediate assessment improved their subject knowledge as shown in the following quotation:

**Table 2.** Student evaluation of teaching results for psychology courses divided in groups.

Course (N <sup>R</sup> ; N <sup>IBP</sup> )	Assignments		Workload		Course	
	Regular	IBP	Regular	IBP	Regular	IBP
<b>Group 1</b>						
Psychology 1 (523; 91)			3.6 (0.8)	4.0 (0.8)	7.6 (1.0)	8.2 (1.1)
Psychology 4 (449; 71)	3.2 (1.1)	3.1 (1.1)	3.7 (0.7)	3.8 (0.8)	6.8 (1.2)	7.2 (1.2)
Psychology 11 (458; 23)	4.0 (1.0)	4.0 (0.9)	3.3 (0.6)	3.4 (0.6)	6.9 (1.4)	6.8 (0.8)
<b>Group 2</b>						
Psychology 3 (495; 95)	2.5 (1.1)	3.3 (1.1)	3.5 (0.8)	3.5 (0.7)	4.8 (1.8)	7.4 (1.4)
Psychology 5 (451; 72)	3.5 (1.2)	3.7 (1.2)	3.6 (0.7)	3.4 (0.7)	7.4 (1.2)	8.1 (1.2)
Psychology 6 (443; 79)	3.1 (1.2)	3.8 (1.0)	3.3 (0.7)	3.5 (0.7)	7.3 (1.0)	7.9 (1.2)
Psychology 7 (485; 72)			3.4 (0.6)	3.4 (0.8)	6.4 (1.4)	5.7 (1.9)
Psychology 10 (427; 63)	3.2 (1.0)	3.8 (0.8)	3.5 (0.7)	3.6 (0.8)	6.5 (1.3)	6.6 (1.3)
<b>Group 4</b>						
Psychology 2 (506; 91)			3.1 (0.6)	3.2 (0.7)	7.7 (0.9)	7.9 (1.1)
Psychology 8 (506; 84)			3.6 (0.7)	3.6 (0.7)	7.1 (1.4)	7.6 (1.4)
Psychology 9 (449; 72)			3.4 (0.6)	3.5 (0.7)	7.4 (1.0)	7.3 (1.6)

Note: N<sup>R</sup> = response rate in the regular bachelor, N<sup>IBP</sup> = response rate in IBP.

**Table 3.** Student evaluation of teaching results for law and criminology courses divided in groups.

Course (N)	Course	
	Law	Criminology
<b>Group 1</b>		
Law 3 (255)	3.5 (0.8) <sup>a</sup>	3.5 (0.8) <sup>a</sup>
Law 10 (97)		3.8 (0.6)
Law 11 (105)		3.9 (0.5)
<b>Group 2</b>		
Law 2 (124)		4.1 (0.5)
Law 4 (126)		3.7 (0.7)
Law 5 (150)	3.7 (0.8) <sup>a</sup>	3.7 (0.8) <sup>a</sup>
Law 7	n.a.	
Law 8 (147)	3.8 (0.7) <sup>a</sup>	3.8 (0.7) <sup>a</sup>
Law 9 (92)		2.9 (0.7)
<b>Group 3</b>		
Law 1 (200)	3.9 (0.7) <sup>a</sup>	
Law 6 (122;117)		3.8 (0.7); 3.8 (0.8)

Notes: No student evaluations of teaching were administered for the Law 7 course in the 2013–2014 academic year. Law 6 was a two-part course and both parts were evaluated separately.

Courses Law 3, Law 5, and Law 8 are part of the Law programme as well as the Criminology programme, and therefore students from both bachelor programmes answer the same SET within these courses.

<sup>a</sup>Denotes online SETs, all others are hard copy.

It kind of made you want to do the material and you really learned it before and then you could read it again afterwards and then it was kind of stuck in your brain and that was good. (IBP student 1)

Third, students mentioned that they liked the fact that the intermediate assessment lowered the stakes of the final exam, since their course grade was no longer totally dependent on their final exam results.

One conceptual criticism students mentioned was that they preferred intermediate assessments that had a clear connection with the final exam. The following quotation illustrates this:

Those essays were pretty annoying because they had something to do with the book, but not really. And it was nothing not really about the exam or about the lectures. (IBP student 2)

They did not endorse the problems with exam training that some teachers reported, that is, no students mentioned that there could be such a thing as too much practice.

Like the teachers, students mentioned that a major practical problem with intermediate assessment was the workload. Most of them mentioned it being heavy, especially when they were just starting their studies. This finding aligns with the SET scores on workload as shown Table 2 (with all mean scores above 3.0):

This was massive in the beginning [...] you try to really know them by heart, [...] and it's insane in the beginning and I was like the walking dead after this. (IBP student 3)

Some students did, however, also mention the workload seeming lighter for courses that they enjoyed and the work being "tough but useful" (IBP student 4). One major problem that is related to the workload was competition with other study activities. A few students mentioned rushing through the course work because they did not have enough time, or deciding not to complete non-obligatory work, like some readings and homework, because they were spending all their time on the intermediate assessments.

Students' detailed accounts on the SET indicated that Law students felt that the intermediate assessment was a good way to familiarise themselves with the demands of a course. One student commented that "The partial exam is a useful tool to see whether you understood the first part correctly". This desire to use the assessment to practise came up in several detailed accounts in the SET. Furthermore, several comments suggest that students wanted more personalised feedback, also indicating that they were looking for pointers on how to fulfil their assignments correctly.

Like the Psychology students, the Law students' main conceptual criticism was that intermediate assessment was often unrelated to the final exam. One comment was that "the paper had little to do with the material".

On the more practical side, the Law students rarely complained about the workload of the different assessments. However, in one course students protested that they had to hand in their assessment at the beginning of the week, before they had the opportunity to follow lectures and workgroups and enhance their understanding of the material.

## Discussion and conclusion

In this research, we investigated how intermediate assessment is perceived by teachers and students. Their perceptions were clustered into (1) general positive opinions, which mainly concerned the fact that students kept up with study work; (2) conceptual criticisms, where teachers and students voiced their critiques on the preconditions and content of the intermediate assessments, and (3) practical considerations, where the main issue discussed was workload.

One positive finding of the current study was that both teachers and students value intermediate assessment because it guided study behaviour and made students study more. Some students mentioned that they would not have kept up with the coursework had the intermediate assessment not been there. Students in the paper by Harland et al. (2015) also feared that they would not make effective use of independent study time and therefore appreciated frequent assessments to serve as a guide for studying.

The relation between student time-use and student results has been extensively studied (e.g. Admiraal, Wubbels, and Pilot 1999; Doumen, Broeckmans, and Masui 2014; Michaels and Miethe 1989). However, the results have not been straightforward. Doumen, Broeckmans,

and Masui (2014) for example, found that the influence of time use on course grade was course dependent, whereas Michaels and Miethe (1989) observed that, for the undergraduate students in their sample, study time correlated positively with results for students who studied during the week, but not for students who crammed at the last minute. Admiraal, Wubbels, and Pilot (1999) concluded that increased time on task led to better study results for their sample.

In the current situation, intermediate assessments could be seen as an incentive to study, but several teachers argued that university students are adults and should take ownership of their study process and their learning. Harland et al. (2015) raise the question of who is responsible for student learning and refer to the system of frequent assessment as a pedagogy of control. With regard to the question of responsibility, van Berkel, Jansen, and Bax (2012) note that universities should take measures to optimise the feasibility of their programmes and that student responsibility comes into play only within these optimised programmes. One measure they suggested to increase feasibility is encouraging students to study by assessing them regularly (Cohen-Schotanus 1999), which is the measure taken by the programmes under investigation.

The favourable attitudes towards intermediate assessment are, however, characterised by one remarkable difference in opinion between teachers and students. Teachers felt that intermediate assessments allowed them to test a broad range of knowledge and skills, but students preferred assessments that measured similar knowledge and skills to those in the final exam. This student preference for overlap is partially supported by testing effect studies that indicate that an overlap in questions is necessary for favourable final exam results (Kling et al. 2005; McDaniel, Wildman, and Anderson 2012). Interestingly, even though students reported that being tested on the same content twice improved their learning, survey studies with American college students (Karpicke, Butler, and Roediger 2009) and Dutch secondary school students (Dirkx 2014) have shown that students often do not use self-testing as a strategy while studying. Apparently, students need to be prompted to perform this kind of intermediate testing.

The difference in preference for content overlap between teachers and students can possibly be explained by the fact that teachers and students have different goals. As mentioned above, teachers' goals were to measure a broad set of knowledge and skills, whereas students' goals were to pass the exam. A few teachers voiced their concerns that the intermediate assessments could lead to exam training and teaching to the test. Teaching to the test is a type of teaching that focuses on passing the final exam, instead of on furthering the general knowledge of the curriculum content (Popham 2001). To avoid teaching to the test our teachers' assessment goals were more focused on assessing different skills and knowledge, instead of simply improving the knowledge of one specific subdomain. Jessop, El Hakim, and Gibbs (2014) noted that assessment goals are often not clear to students, which might also have been a factor in the current discrepancy in preference.

A more practical issue that both teachers and students faced was the workload that is associated with intermediate assessment. For students this workload seemed to lead to competition between the different course activities. Non-mandatory study activities got pushed aside for intermediate assessments, especially if they were graded. Vos (1992) had already suggested that competition may be an issue for students, if programmes do not keep student workload in mind while designing their curriculum. In their interview study, Harland et al. (2015) also encountered strong evidence for competition, where students



walked out of a non-graded lecture because they would rather spend time on their assessment. One possible solution to this problem is to take a more programmatic approach towards assessment (van der Vleuten et al. 2012). Programmatic assessment involves a more comprehensive approach, where several low-stake formative assessments build up to a high-stake summative decision. When the assessment programme is designed as a whole instead of as different components, teachers are more aware of the possible workload faced by students.

The fact that teachers mentioned their workload is in concordance with the results of Goos, Gannaway, and Hughes (2011), who found that teachers' main concerns were regarding assessment conditions. With regard to teacher workload, it is important to keep the assessment goals in mind. Harland et al. (2015) encountered the problem that teachers felt there was no time to provide feedback because of the large number of assessments. Leeming (2002) also noted that frequent assessment comes with a sizeable workload. However, he suggests adapting the type of assessment to reduce workload; for example, by reducing the number of essay questions instead of the number of assessments.

### **Limitations**

The sample of teachers, courses and assessment types is relatively small and selected from four programmes in one university in the Netherlands. Furthermore, all teachers that were interviewed during this study were part of a department-wide educational innovation project. They were actively engaged in improving their courses during this project. The fact that the teachers were participating in an innovation project could have had an influence on their motivation and their perceptions of intermediate assessment. Teachers in the context of Law, Criminology and Psychology programmes undergoing educational innovation projects are only a small facet of the total population of university teachers and therefore generalising to the general population of university teachers is inadvisable. We do think our findings have generated plausible hypotheses to be tested in future research. Future research could expand the research population in two directions. First, they could include teachers from different departments and other universities. Second, teachers who are not a part of an educational innovation should be investigated about their perceptions.

### **Implications for practice**

The limited sample size in the current papers precludes bold implications, but we believe there are three general indications for practice that can follow from our results and the literature discussed in this section. First, students need to be made more aware of the goals of intermediate assessment, as better awareness would probably add to their positive perception of it. Second, teachers and students need to be made explicitly aware of the cognitive benefits of intermediate assessment on top of behavioural effects, like keeping on track. The third implication is that measures should be taken to alleviate teacher and student workload. One of these measures could be to take a programmatic approach towards assessment, to spread out study activities and reduce competition between parallel courses.

## Concluding remarks

University education is subject to change and curriculum innovations often follow each other in quick succession. These innovations are usually evaluated by looking at grades, pass percentages, or by using questionnaires. However, teachers and students need to support the innovation for any of these measures to be useful. The results from this study suggest that teachers and students at the four programmes under investigation generally have positive perceptions of intermediate assessment, along with some conceptual criticisms and with concerns about practical considerations. The assessment type of a course does not seem to play a role in these perceptions, but student and teacher characteristics might. Future research should take these characteristics into account. Research by Baeten, Dochy, and Struyven (2012) has already indicated that student perceptions of their learning environment are influenced by characteristics such as their motivation.

## Note

1. Quotes have been translated from Dutch and edited for length and legibility where applicable.

## Acknowledgements

The authors would like to thank all teachers and students who participated in the interviews.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

**Indira Day** is a PhD candidate at the Leiden University Graduate School of Teaching. Her research focuses on the relation between intermediate assessment, study behaviour and study success in higher education.

**Floris van Blankenstein** is an educational developer and researcher at the Leiden University Graduate School of Teaching. His research focuses on blended learning, collaborative learning and peer feedback in higher education.

**Michiel Westenberg** is a full professor of developmental psychology and scientific director of the Institute of Psychology at Leiden University. His research interests include motivation and fear of failure in high school populations, specifically fear of negative evaluation in performance and social interaction situations.

**Wilfried Admiraal** is a full professor of educational sciences and director of the Leiden University Graduate School of Teaching. His research combines education, technology and social psychology in secondary and higher education.

## ORCID

Indira N. Z. Day  <http://orcid.org/0000-0002-5511-1061>  
W. F. Admiraal  <http://orcid.org/0000-0002-1627-3420>

## References

- Admiraal, W., T. Wubbels, and A. Pilot. 1999. "College Teaching in Legal Education: Teaching Method, Students' Time-on-Task, and Achievement." *Research in Higher Education* 40 (6): 687–704.
- Baeten, M., F. Dochy, and K. Struyven. 2012. "Using Students' Motivational and Learning Profiles in Investigating Their Perceptions and Achievement in Case-Based and Lecture-Based Learning Environments." *Educational Studies* 38 (5): 491–506.
- Bennett, R. E. 2011. "Formative Assessment: A Critical Review." *Assessment in Education: Principles, Policy & Practice* 18 (1): 5–25. doi:10.1080/0969594X.2010.513678.
- van Berkel, H., E. Jansen, and A. Bax, eds. 2012. *Studiesucces Bevorderen: het kan en het is niet moeilijk* [Improving Study Success: It's Possible and It's Not Difficult]. The Hague: Boom Lemma.
- Bongers, V. 2015. "Betutteling ligt op de loer: Studenten klagen over hoeveelheid Toetsen." [Paternalism Looms: Students Complain about the Amount of Tests.] *Mare*, February 19. <http://mareonline.nl>.
- Bouwmeester, R. A. M., R. A. M. de Kleijn, A. W. M. Freriksen, M. G. van Emst, R. J. Veeneklaas, M. J. W. van Hoeij, M. Spinder, M. J. Ritzen, O. T. J. Ten Cate, and H. V. M. van Rijen. 2013. "Online Formative Tests Linked to Microlectures Improving Academic Achievement." *Medical Teacher* 35 (12): 1044–1046.
- Cohen-Schotanus, J. 1999. "Student Assessment and Examination Rules." *Medical Teacher* 21 (3): 318–321. doi:10.1080/01421599979626.
- van Dinther, M., F. Dochy, M. Segers, and J. Braeken. 2014. "Student Perceptions of Assessment and Student Self-Efficacy in Competence Based Education." *Educational Studies* 40 (3): 330–351.
- Dirkx, K. J. H. 2014. *Putting the Testing-Effect to the Test: Why and When is Testing Effective for Learning in Secondary School?* The Netherlands: Datawyse, Universitaire Pers Maastricht.
- Doumen, S., J. Broeckmans, and C. Masui. 2014. "The Role of Self-Study Time in Freshmen's Achievement." *Educational Psychology* 34: 385–402. doi:10.1080/01443410.2013.785063.
- Dunlosky, J., K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham. 2013. "Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology." *Psychological Science in the Public Interest* 14: 4–58. doi:10.1177/1529100612453266.
- Entwistle, N. J. 1991. "Approaches to Learning and Perceptions of the Learning Environment." *Higher Education* 22 (3): 201–204.
- Gibbs, G., and C. Simpson. 2004. "Conditions under Which Assessment Supports Students' Learning." *Learning and Teaching in Higher Education* 1 (1): 3–31.
- Goos, M., D. Gannaway, and C. Hughes. 2011. "Assessment as an Equity in Higher Education: Comparing the Perceptions of First Year Students, Course Coordinators, and Academic Leaders." *The Australian Educational Researcher* 38: 95–107.
- Harland, T., A. McLean, R. Wass, E. Miller, and K. N. Sim. 2015. "An Assessment Arms Race and Its Fallout: High-Stakes Grading and the Case for Slow Scholarship." *Assessment & Evaluation in Higher Education* 40: 528–541. doi:10.1080/02602938.2014.931927.
- Jessop, T., Y. El Hakim, and G. Gibbs. 2014. "The Whole is Greater than the Sum of its Parts: A Large-Scale Study of Students' Learning in Response to Different Programme Assessment Patterns." *Assessment & Evaluation in Higher Education* 39: 73–88. doi:10.1080/02602938.2013.792108.
- Karpicke, J. D., A. C. Butler, and H. L. Roediger III. 2009. "Metacognitive Strategies in Student Learning: Do Students Practise Retrieval When They Study on Their Own?" *Memory* 17: 471–479. doi:10.1080/09658210802647009.
- Kling, N., D. McCorkle, C. Miller, and J. Reardon. 2005. "The Impact of Testing Frequency on Student Performance in a Marketing Course." *Journal of Education for Business* 81 (2): 67–72.
- Kornell, N. 2009. "Optimising Learning Using Flashcards: Spacing is More Effective than Cramming." *Applied Cognitive Psychology* 23 (9): 1297–1317. doi:10.1002/acp.1537.
- Leeming, F. C. 2002. "The Exam-a-Day Procedure Improves Performance in Psychology Classes." *Teaching of Psychology* 29: 210–212.
- Lizzio, A., K. Wilson, and R. Simons. 2002. "University Students' Perceptions of the Learning Environment and Academic Outcomes: Implications for Theory and Practice." *Studies in Higher Education* 27 (1): 27–52.
- McDaniel, M. A., K. M. Wildman, and J. L. Anderson. 2012. "Using Quizzes to Enhance Summative-Assessment Performance in a Web-Based Class: An Experimental Study." *Journal of Applied Research in Memory and Cognition* 1: 18–26. doi:10.1016/j.jarmac.2011.10.001.

- Michaels, J. W., and T. D. Miethe. 1989. "Academic Efforts and College Grades." *Social Forces* 68: 309–319.
- Myers, C. B., and S. M. Myers. 2015. "The Use of Learner-Centered Assessment Practices in the United States: The Influence of Individual and Institutional Contexts." *Studies in Higher Education* 40 (10): 1904–1918.
- Popham, W. J. 2001. "Teaching to the Test?" *Educational Leadership* 58: 16–21.
- Prosser, M., P. Walker, and R. Millar. 1996. "Differences in Students' Perceptions of Learning Physics." *Physics Education* 31 (1): 43–48.
- Roediger III, H. L., and J. D. Karpicke. 2006. "The Power of Testing Memory: Basic Research and Implications for Educational Practice." *Perspectives on Psychological Science* 1 (3): 181–210. doi:[10.1111/j.1745-6916.2006.00012.x](https://doi.org/10.1111/j.1745-6916.2006.00012.x).
- van der Vleuten, C. P. M., L. W. T. Schuwirth, E. W. Driessen, J. Dijkstra, D. Tigelaar, L. K. J. Baartman, and J. van Tartwijk. 2012. "A Model for Programmatic Assessment Fit for Purpose." *Medical Teacher* 34 (3): 205–214.
- Vos, P. 1992. "Het ritme van het rooster." [The Rythm of the Schedule.] *Onderzoek van Onderwijs* 21: 51–53.