The following handle holds various files of this Leiden University dissertation:
http://hdl.handle.net/1887/63156

**Author:** Day, I.N.Z.
**Title:** Intermediate assessment in higher education
**Issue Date:** 2018-06-28
Chapter 6

General discussion
6.1 Introduction

In this dissertation I investigated intermediate assessment characteristics, concentrating on three focal points (implementation, perceptions, and effects). Chapter 2 is an overview of the literature, whereas Chapters 3 – 5 focus on the introduction and subsequent use of intermediate assessment in the curricula of four programs at Leiden University. To reiterate, intermediate assessment is defined as all assessment that takes place during the course period, instead of at the end. In Chapters 3 – 5, an added requirement for intermediate assessment was that the assessment was mandatory and needed to be handed in to the teacher, to prevent self-selection biases that are associated with voluntary assessments.

6.2 Main findings per chapter

Chapter 2. A Review of the Characteristics of Intermediate Assessment and their Relationship with Student Grades

This chapter provides an overview of the literature on intermediate assessments in higher education, with a specific focus on several assessment characteristics. The research question central to this chapter is ‘what characteristics of intermediate assessment are related to student grades.’ In total 88 articles using various types of intermediate assessment were synthesised. These articles showcased the variety of intermediate assessment in higher education: 24 different assessment types were used across many different disciplines. There was also a large variation in frequency of assessment, ranging from a single assessment to 38 assessments. Four main characteristics were further discussed. These were use of feedback, mandatory assessments, the assessor, and the rewards for the assessment. In general, the majority of studies found positive effects of intermediate assessment on student grades. When looking at the four main characteristics, it seems that studies using corrective feedback found positive influences on student grades most often, but literature (Shute, 2008) suggests that elaborate feedback may be more beneficial for some groups. There does not seem to be a definitive answer whether an assessment should be mandatory, because a similar proportion of articles using mandatory assessment as articles using voluntary assessments found positive results for student grades. Peer and self-assessment showed promising results in the studies employing it, but it should be kept in mind that
not all assessment is suitable for peer or self-assessment, and furthermore the number of articles using teacher assessment largely outnumbered those using peer or self-assessment. With regard to the reward for assessment, studies providing students with course credit as a reward for completing intermediate assessment showed the most positive results. This chapter ends with three scenarios for intermediate assessment, to further exemplify how intermediate assessment can be deployed in higher education.

Chapter 3. Teacher and Student Perceptions of Intermediate Assessment in Higher Education

Chapter 3 is a first exploration of the perceptions of intermediate assessment and their types. The study described in this chapter was conducted in the first year of the educational reform introducing intermediate assessment. Teachers involved in teaching first-year courses using intermediate assessment were interviewed and information from student evaluations of teaching and student interviews was used to gauge these perceptions. The research questions for this chapter are:

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?

Contrary to the expectations, there were no differences in perceptions for different assessment types. Overall it can be said that teachers were positive about intermediate assessment as a way to get students to keep up with study work, and as a possibility to measure a variety of knowledge and skills. However, they were not all in agreement with the obligatory nature of intermediate assessment, and some felt that the fact that students could compensate their intermediate assessment grade was not desirable. Furthermore, teachers remarked a few practical concerns, focusing on the added workload of correcting assessments, and administration that came with the assessment. Other practical concerns were relating to possible plagiarism and resits of the assessment. Students were also generally positive about intermediate assessment, but for different reasons. The main attraction for students seemed to be that intermediate assessment kept them on track, that it deepened their knowledge, and that it lowered the stakes of the final exam. In complete opposition to teachers’ appreciation for measuring a variety of knowledge and skills, the main critique of students was that sometimes
the intermediate assessment was not related to the final exam. Furthermore, students also complained about the workload of the intermediate assessments.

**Chapter 4. University Teachers’ Conceptions of their Current and Ideal Intermediate Assessment**

Following the exploration of perceptions in Chapter 3, Chapter 4 delves deeper into teachers’ ideas about assessment. To research the question ‘What differences in conceptions of intermediate assessment do university teachers display when discussing their current and ideal intermediate assessment?’ thirteen course coordinators of first-year Law, Criminology, and Psychology courses were interviewed on their ideas about their current assessment and how they would employ intermediate assessment in an ideal situation. Teachers’ responses were coded using codes from two existing sources (Postareff Virtanen, Katajavuori, and Lindblom-Ylänne, 2012; Samuelowicz and Bain, 2002). In both these sources teachers’ ideas can be identified as either reproductive (focusing on memorisation and reproduction of knowledge) or transformational (focusing on knowledge construction). Overall it could be said that teachers had transformational goals for their current as well as their ideal assessment, but goals for the ideal assessment were more transformational. Teachers were divided into subgroups based on the differences in conceptions for the current and ideal assessment. One teacher had no transformational goals for both, three teachers had a mix of transformational and reproductive goals for both, and one teacher had only transformational goals. Furthermore, seven teachers had more transformational goals for their ideal assessment, and one teacher had more transformational goals for her current assessment.

**Chapter 5. Explaining Individual Student Success Using Intermediate Assessment Types and Student Characteristics**

In this chapter assessment type was used to quantitatively investigate the relationship between assessment characteristics and student grades. Furthermore, student characteristics were added to this equation, to get a fully informed vision of the subject. The two research questions central to this chapter were:

1. To what extent does the type of intermediate assessment relate to academic achievement?
2. What role do gender, high school achievement, motivation and self-regulation play in this relationship?
To investigate this, an adapted version of the motivated strategies for learning questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993) was filled in by 94 first-year law and criminology students. Their final course grades were accessed and courses were divided in six assessment type groups (no intermediate assessment, written assignment, homework assignments, partial exam, interview and paper, presentation propositions and mini experiment). The final two type groups were excluded from analysis, because less than ten students in the sample completed the course. After initial regression analyses indicated which variables were related to outcomes, repeated measures ANCOVA's were conducted to discover whether assessment type and student characteristics influenced student grades. Results indicate that assessment type does not influence students’ grades and high school GPA is the strongest predictor of achievement. Furthermore, male students scored worse on courses without intermediate assessment, but this gender achievement gap disappeared in courses using any type of intermediate assessment. Lastly, two measures for intrinsic motivation (i.e., task value and intrinsic goal orientation) were negative predictors for two types of intermediate assessment, suggesting that some types of intermediate assessment may be detrimental for students with higher levels of intrinsic motivation.

6.3 General findings

The overview of the literature in Chapter 2 shows the wide variety of ways in which intermediate assessment can be implemented in the curriculum. A closer look at the implementation at Leiden University in Chapters 3-5 also shows several different types of assessment characteristics. Furthermore, teachers’ ideas about their ideal assessment in Chapter 4 show that there could be even more different assessment types, teachers had full control of the educational environment and its circumstances.

With regard to the perceptions of intermediate assessment, chapters 3 and 4 show that teachers generally have positive perceptions. Chapter 3 shows that teachers appreciate that intermediate assessment can help keep students on track and that they can measure a variety of knowledge and skills. Additionally, Chapter 4 shows that teachers are very focused on using intermediate assessment to develop student learning, and that they would focus on student learning even more in their ideal situation.
The effects of intermediate assessment show a less unequivocal picture. The literature review in Chapter 2 indicates that intermediate assessment generally has a positive effect on students’ grades, but the results from the quantitative study described in Chapter 5 do not show any influence of assessment. In this study students’ high school GPA was the main predictor of university grades. Whether a course had intermediate assessment or not, or what type of assessment this was, if there was one, did not seem to matter.

6.4 Reflections on the design of intermediate assessment

Since this dissertation is specifically focused on the characteristics of intermediate assessment, the following section will go deeper into the perceptions and results specifically related to two important assessment characteristics.

6.4.1 Assessment Type

A main finding is that the type of intermediate assessment is not a defining characteristic. This is evidenced in the literature review (Chapter 2) as well as in two of the empirical studies (Chapters 3 and 5). Both the review and Chapter 5 indicate that the assessment type does not influence student grades. In the review it is evidenced that for the majority of assessment types there are more studies showing positive effects of intermediate assessment than there are studies that do not portray these effects. There are some assessment types where this finding is reversed (like cases and exercises) and a few where there are only negative results. However, these are usually small groups of five or less studies. For the three largest assessment type groups, ‘general assessment’, quiz, and writing assignment, there is a similar ratio where two-thirds of the studies have found positive results. In Chapter 5, the lack of influence of assessment type is shown by the fact that the type of assessment is not a significant factor in the repeated measures ANCOVA, indicating that students get similar final exam scores regardless of the type of intermediate assessment employed in their course. Chapter 3 shows that this lack of influence of assessment type reaches further than just student results. Teachers across courses using different types of intermediate assessment all had similar perceptions about intermediate assessment. This indicates that, at least for assessment type, it does not really matter how you employ intermediate assessment. Results from Chapter 5 even
seem to suggest that it does not matter whether you use intermediate assessment, contrasting with results from previous research (e.g., Larsen, Butler, & Roediger, 2008; McDaniel, Anderson, Derbish, & Morriette, 2007).

A possible explanation for the current lack of an effect for (different types of) intermediate assessment, that was also raised in Chapter 5, is that all types of intermediate assessment promote distributed studying. Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013) have shown that distributed practice improves student results. Furthermore, distributed studying may also lead to increased study time, which in turn may also lead to better results (Doumen, Broeckmans, & Masui, 2014). However, since university curricula are usually set up with distributed educational meetings that have required readings, it stands to reason that the distributed practice effect is enhancing results in all university courses, not just those using intermediate assessment.

Another explanation could be that students did not perceive the intermediate assessment as aligned with, and therefore useful for, the final exam. Kahu (2013) proposes that assessment can influence students’ engagement, and subsequently their study behaviour and results. Students’ preference for a strong connection between the intermediate and final assessment already came up in Chapter 3, and it stands to reason that students are less engaged if they perceive this connection to be absent.

### 6.4.2 Use of Feedback

With regard to other characteristics, the importance of feedback is underlined. Chapter 2 shows that almost all studies using intermediate assessment provided some form of feedback, and that corrective feedback seemed the most positive. When addressing their beliefs about assessment in Chapter 4, teachers focused on student learning and the importance of providing feedback in this process.

Feedback is a construct that is thoroughly investigated, whether in the case of formative assessment (Nicol & Macfarlane-Dick, 2006), or when referring to learning and education in general (Hattie & Timperley, 2007). Gibbs and Simpson (2004) mention ten conditions that assessment has to meet to improve learning, seven of which focus on the influence of feedback. The recommendations for instructive feedback given by different authors often overlap. Several authors, for example, suggest providing task or process focused feedback (Gibbs & Simpson, 2004; Hattie & Timperley, 2007; Shute, 2008). The finding that corrective feedback may be sufficient is partly corroborated by
Shute (2008) who suggests that feedback should be simple but specific enough. Other suggestions include that feedback should inform students on how they can attain the desired performance (Nicol & Macfarlane-Dick, 2006; Hattie & Timperley, 2007, refer to this as feed-forward), and that teachers should be able to reshape their teaching with feedback information (Nicol & Macfarlane-Dick, 2006). Several teachers in Chapter 4 also referenced how feedback could help them improve their teaching.

6.5 Limitations

None of the chapters of this dissertation have established base-line measures, or control groups to compare results to. Establishing a baseline was not possible, because the educational reform was studied in real time during its roll-out. Similarly, in educational research, using true control groups is often difficult, because students are all offered the same curriculum.

A second limitation for all four studies described in Chapters 2 to 5 is issues with selection. Teachers and students could self-select for participation in the interviews and questionnaires. It is possible that teachers or students who felt very negative about the concept of intermediate assessment declined to participate, and therefore their opinions are not included in this dissertation. However, for both interview studies (Chapters 3 and 4) the teacher participation rate is approximately 70 percent, indicating that even if teachers who were very negative about intermediate assessment chose not to participate, the positive results portrayed in this dissertation do represent the majority of teachers. Student participation rates (Chapter 5) are much smaller, but the results show that participating students displayed a range of final exam grades. Similarly, the literature review in Chapter 2 suffers from its own selection issue: publication bias. In general, articles displaying positive results are more likely to be accepted for publication in scientific journals. Mahoney (1977), for example, already showed that reviewers deem articles with positive outcomes to be more methodologically sound. Shadish, Doherty, and Montgomery (1989) estimated that there may be as many unpublished studies in their field as there are published ones. Since the review did not include unpublished articles, we cannot be sure that the full spectrum of intermediate assessment characteristics in higher education is represented in the review.
In addition to the problems with selection, the three empirical studies all describe relatively small samples of teachers or students from a limited number of educational programs at Leiden University. These programs are all large scale, with few contact hours. It stands to reason that these results may not generalise to programs with a smaller number of students or more contact hours. Several STEM programs, for example, have lectures in the morning and practical labs in the afternoon, eliminating at least the need to introduce intermediate assessment as a means to get students to work. However, the review in Chapter 2 does describe studies from different educational fields, which often have positive results.

6.6 Implications for Practice

6.6.1 Encourage use of intermediate assessment
In general, intermediate assessment seems to be beneficial for students’ performance. The overview of the literature presented in Chapter 2 evidences this. Teachers and students presented several benefits of intermediate assessment in Chapter 3, including the ability to measure different skills, and to take pressure of the final exam. Results from previous studies also show benefits when looking at affective or behavioural outcomes, intermediate assessment can, for example, be used to engage students (Holmes, 2015) or encourage them to spend more time studying (Admiraal, Wubbels & Pilot, 1999).

6.6.2 Minimise institutional assessment guidelines
The studies described in this dissertation suggest that a multitude of different assessment characteristics can influence student results. Therefore, it seems important to keep requirements for intermediate assessment as minimal as possible, to facilitate teachers’ independence in deciding on the type and characteristics of the assessment they want to provide. When course coordinators were interviewed by Goos, Gannaway and Hughes (2011), one of their main problems with assessment was ‘increasingly centralised and bureaucratic assessment requirements’ (p. 99). The current results seem to advise against constraining assessment practice too firmly.
6.6.3 Support teachers’ assessment practice

A main conclusion from Chapters 3 and 4 is that teachers are willing to implement intermediate assessment. Often, the teachers have positive perceptions and ambitious ideas on how they would like to incorporate intermediate assessment in the courses they are teaching. However, to keep this enthusiasm intact, it is important that the assessment practice is supported. Goos et al. (2011) identified that, on top of the problems with bureaucratic requirements discussed in the previous section, teachers’ main problem with regard to assessment was related to workload and lack of tailored support.

Workload was identified as a constraining factor by teachers quoted in this dissertation as well. Several teachers mentioned, for example, the heavy workload of providing students with individual feedback. Most university courses are taught by staff whose primary task is research, creating a tension between research and teaching. When an institution wants to fully utilise the advantages of intermediate assessment, it should be prepared to invest in measures to alleviate workload. A utopian solution would be to increase the amount of teaching staff, but more realistically, institutions can invest in (digital) tools to support teachers’ assessment practice. Several tools exist that can alleviate the burden of workload, for example by easing the feedback process.

Additionally, it is important to support teachers’ knowledge about assessment. Postareff et al. (2012) suggested that teachers need to be made aware of the benefits of transformational assessment. Chapter 4 suggests that most teachers in the sample are already aware of these benefits, but nonetheless, assessment knowledge could be further supported. Some teachers quoted in Chapter 3 mentioned being unsure about what type of assessment was the best fit for their course, indicating that their assessment knowledge could be broadened. Samuelowicz and Bain (2002) mention that there is a large chance that several teachers enrolled in staff development classes will hold widely different beliefs on assessment, which may mean that a one size fits all assessment knowledge course will not work. A potent way to have teachers exchange their assessment knowledge is by using communities of practice, or another group-like constellation where teachers can learn from each other’s experiences.
6.7 Directions for future research

Future studies into the characteristics of intermediate assessment can expand on this dissertation in several ways. First of all, this research can focus more quantitatively on the behavioural and affective outcomes of intermediate assessment. It could, for example, take a similar approach as in Chapter 5, but to measure more whether specific types of intermediate assessment elicit more student studying behaviour than others do.

A second direction for future research could focus further on the assessment types. The majority of studies cited in Chapter 2, and the studies conducted in Chapters 3 to 5, focused on comparing the effects of assessment types between courses. Perdigones, Garcia, Valino, & Raposo (2009), for example, investigated what they called comparable courses and found that a system of writing an article for a professional journal and presenting this article elicited higher pass rates for their course than a system of continuous assessment did. However, these ‘comparable’ courses differed in the number of credits, the university they were taught at, and the number of students enrolled. Furthermore, it stands to reason that several other course characteristics may differ between courses as well, thereby influencing the outcomes. Future studies could focus on comparing different assessment types within a course, either in a cohort design (keeping everything constant between cohorts, except for the assessment) or by comparing several groups in a quasi-experiment, where students are either assigned to a different assessment type or options for intermediate assessment are given, and students choose their preferred assessment type.

Results from Chapter 5 indicate that several student characteristics may influence the relationship between intermediate assessment and student results. Teachers could choose to offer tailored assessment for different student groups, and subsequent research could investigate whether this tailored assessment leads to improved performance for these groups. Moreover, teachers could offer several intermediate assessment options and let students choose their own assessment program. In this case, it is interesting to study whether different groups of students choose different types of assessment.

Finally, future research could further explore the influence of student characteristics. Richardson, Abrahams, and Bond (2012) identified 41 predictors of student achievement in their meta-analysis, of which only ten were discussed in the current dissertation. It stands to reason that several of the other
characteristics are also related to intermediate assessment and student outcomes. Future research could focus on (self-regulated) learning strategies (which are somewhat related to the motivation and self-regulation discussed in Chapter 5), or on personality traits, in a similar design as Chapter 5 to see whether these characteristics predict student results.

6.8 Concluding Remarks

Overall, it can be said that various types of assessment all have positive influences on student results, and that maybe the act of providing students with an intermediate assessment is more important than what the exact assessment is. University teachers often have thorough ideas and ideals about intermediate assessment, but often lack the means to put these into practice. Institutions should support teachers with assessment knowledge courses or opportunities to exchange assessment knowledge with colleagues. Additionally, they should minimise assessment constraints and simultaneously provide teachers with the freedom to design assessments that fit the courses that are being taught.