

A review of the characteristics of intermediate assessment and their relationship with student grades

Indira N. Z. Day, Floris M. van Blankenstein, Michiel Westenberg & Wilfried Admiraal

To cite this article: Indira N. Z. Day, Floris M. van Blankenstein, Michiel Westenberg & Wilfried Admiraal (2018) A review of the characteristics of intermediate assessment and their relationship with student grades, *Assessment & Evaluation in Higher Education*, 43:6, 908-929, DOI: [10.1080/02602938.2017.1417974](https://doi.org/10.1080/02602938.2017.1417974)

To link to this article: <https://doi.org/10.1080/02602938.2017.1417974>



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



Published online: 27 Dec 2017.



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



Article views: 1622





View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

A review of the characteristics of intermediate assessment and their relationship with student grades

Indira N. Z. Day^a , Floris M. van Blankenstein^a, Michiel Westenberg^b and Wilfried Admiraal^a 

^aICLON, Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands; ^bInstitute of Psychology, Leiden University, Leiden, The Netherlands

ABSTRACT

In-course assessment, such as midterms, quizzes or presentations, is often an integral part of higher education courses. These so-called intermediate assessments influence students' final grades. The current review investigates which characteristics of intermediate assessment relate to these grades. In total, 88 articles were reviewed that examined the relationship between intermediate assessment and student grades. Four main characteristics were identified: the use of feedback, whether the assessment is mandatory, who is the assessor, and the reward students get for participating. Results indicate that corrective feedback leads to the most positive results, but elaborate feedback may benefit lower achieving groups. No difference in results was found for mandatory versus voluntary intermediate assessments. Peer assessment seemed to be beneficial, and rewarding students with course credit improves grades more than other rewards. Three scenarios are presented on how teachers can combine the different characteristics to optimise their intermediate assessment.

KEYWORDS

Intermediate assessment; continuous assessment; student grades; literature review

Introduction

Assessment is a powerful force in education. Dunlosky et al. (2013) posit that practice testing is an effective studying technique, next to distributed practice (the process of studying constantly throughout a study period instead of cramming). Since these two processes enhance student learning in self-study, it stands to reason that distributed assessments should be an integral part of a course, encompassing different assessment moments during the course instead of just a final examination. In higher education, these assessments during the course often take various forms.

In the current paper, all forms of assessment during the course will be referred to as 'intermediate assessment' (synonyms are 'frequent' or 'continuous' assessment, see Isaksson 2008; Rezaei 2015). Intermediate assessment may influence three different types of student outcomes, namely affective, cognitive or behavioural outcomes. Some examples of these outcomes are student perceptions of their learning (Bälter, Enström, and Klingenberg 2013), student engagement (Holmes 2015), student studying behaviour (Admiraal, Wubbels, and Pilot 1999) and student grades (De Paola and Scoppa 2011).

Higher education teachers employ intermediate assessment with several goals in mind. Research by Day et al. (2017) indicates that teachers mainly employ intermediate assessment to motivate students to start studying, and to be able to measure a variety of learning goals, knowledge and skills. However,

CONTACT Indira N. Z. Day  i.n.z.day@iclon.leidenuniv.nl

they often expressed uncertainty about the influences of different assessment characteristics on their goals. Therefore, in the current review we further investigate different assessment characteristics.

Intermediate assessment in higher education is a subject that is widely researched. Assessments can vary on a multitude of characteristics, some of which are more thoroughly researched than others. One characteristic is the type of the assessment. Different results may be gained when using quizzes, essays or presentations. Often the type of an assessment is decided by the learning goals a teacher wants to assess: so called constructive alignment (Biggs 1996).

Other characteristics include the use of feedback, or the use of formative assessments (Bennett 2011). Studies show that formative assessments are potent learning tools (Black and William 2004), and that instructive feedback enhances learning. Other examples of characteristics are the frequency of the assessment, who the assessor is and whether students are graded on the assessment or not. The latter might be important to distinguish as students appear to be focused on grades (Harland et al. 2015).

There is a lot of evidence from studies that intermediate assessment has a general positive effect on students' final grades. However, these studies (e.g. Admiraal, Wubbels, and Pilot 1999; De Paola and Scoppa 2011; Leeming 2002) often focus on one specific type of assessment, not contrasting them with other types.

The reality of higher education often prevents teachers from trying out different types of intermediate assessment within the same instalment of their course. The goal of the current paper is to provide an overview of research that has been conducted into intermediate assessment in higher education, across disciplines, and its relation to student grades. We have decided on grades as a student outcome because in educational research, grades are often used as a proxy for cognitive learning (Richardson, Abraham, and Bond 2012). Assessment drives learning (Clouder et al. 2012, 2), making grades a suitable outcome variable.

With this review, the authors aim to provide a more general overview of what intermediate assessments have led to positive outcomes and to inform higher education professionals who are designing their curriculum. We try to answer the following research question: what characteristics of intermediate assessment are related to student grades?

Methods

Search and Selection

Papers for the review were searched using the 'All databases' search in the Web of Science. In each of the searches one of the following keywords was combined with 'assessment': 'intermediate', 'frequent', 'continuous', 'programmatic', 'in-between', 'formative' and 'summative'.

The different steps of the search process, including which parts of the articles were examined, the inclusion criteria and resulting selected articles are presented in Figure 1. Steps 1 and 2 (search and first selection) were conducted in March 2016, which was also the cut-off date for published articles. The year 2000 was chosen as a start date, as the number of published articles sharply decreases the earlier the date. In each subsequent step of the search process, the inclusion criteria were expanded to select more focused studies.

Ultimately, 88 articles were selected for inclusion in the review. These articles included (quasi) experiments as well as descriptive studies. While reading the methods and results sections all relevant information concerning the intermediate assessment and its results was extracted from the articles. All information extracted from the articles fits in the categories presented in Table 1. Context variables and assessment characteristics were defined while reading the first few articles and discussed amongst authors. Furthermore, at the annual meeting of the Netherlands Educational Research Association, educational researchers and educationalists attending our discussion session (Day et al. 2016) were asked to mark which variables they deemed interesting and useful.

All relevant information from the articles was added to a Microsoft Access database. When an article does not provide information on one of the variables, 'no info' is entered. After entering all information

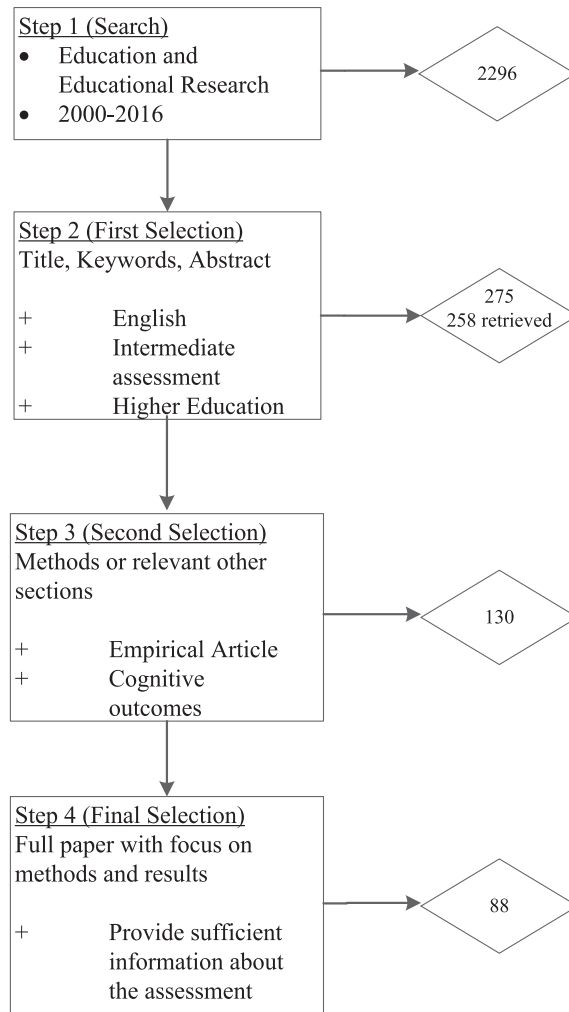


Figure 1. Schematic overview of the search and selection process.

several iterations were performed to make sure similar terms were used for similar concepts across articles.

Reliability of selection

The search and selection process was carried out by the first author. During this process, quality was ensured in two ways. During the search (step 1), the second author selected articles from the ‘continuous’ assessment search in parallel to the first author. Each author made a short overview of why they selected these articles. Subsequently, they discussed and argued their choices till an agreement on which articles to include in step 2 was reached.

Additionally, in step 4, the second author judged a subset of 14 articles from step 3 (a little over 10% of the total) on whether they should be included in the full review. The authors agreed on 12 of the articles, but both included one article that the other did not include. This led to an acceptable inter-rater-reliability of Cohen’s $\kappa = 0.65$. After careful reconsideration of the two articles, one was not relating intermediate assessment to final examination outcomes and was therefore not included. The other article was at first discarded by one author, because the information about intermediate

Table 1. Overview of information extracted from the articles included in the review.

Context variables	Assessment characteristics
Country	Type of assessment e.g. quiz, presentation, essay, cases
Discipline	Frequency of assessment e.g. weekly, single assessment, 8 assessments
N of students	Duration
Study Phase e.g. first year, undergraduate	Number and type of questions e.g. multiple choice, short answer
Use of control group	Scoring e.g. percentage of grade, performance score, bonus points
Cognitive Outcome Variable e.g. final examination score, GPA	Feedback e.g. elaborate, feedback form
Other variables in the research e.g. gender, comparing low and high achieving students	Assessor teacher, peers, self
Findings of the research	Medium e.g. computer, pen and paper
	Location In class, at home, anywhere
	Mandatory
	Other characteristics e.g. open book, reattempts

assessment was limited, but since it did relate intermediate assessment to final examination results it was included in the final selection.

Analysis

Several papers describe research on multiple cohorts, multiple assessment types or multiple studies. During the analyses, these will be treated separately, raising the total number of outcomes over 88.

Relevant characteristics for analysis were decided upon in discussion with educational experts (Day et al. 2016), and inspired by the typology of peer assessment by Topping (1998), as, to our best knowledge, a similar typology of intermediate assessment is not available. Furthermore, characteristics that seemed salient while reading the articles were added as well. All characteristics were investigated while articles were grouped on the assessment type level. Table 2 shows an overview of all articles and their assessment characteristics. The table is organised by assessment type, and each condition (or assessment type) is presented on a separate row, to be able to clarify the differences in characteristics within the same articles.

To further exemplify the clusters of characteristics of intermediate assessment that are present in the literature, illustrative articles for the clusters will be discussed in more detail. These articles are selected because they present relevant results pertaining to the cluster, but they are not supposed to be representative of all studies included in the cluster.

Results & discussion

Article background information

Research into intermediate assessment in higher education is conducted all over the world, evidenced by papers from South Africa (3), Iran (3) and Brazil (1). However, the majority of the research discussed in the current review was conducted in the United States of America (25), Australia (12) and the United Kingdom (12).

Similarly, research is being conducted in different disciplines, with medicine (38), STEM (science, technology, engineering and mathematics) fields (19), the social sciences (12), and law, business and economics (12) being main contributors. The largest group of studies is conducted with first year students (with 23 articles exclusively using first year students).



Table 2. Overview of most important assessment characteristics of the articles included in the review.

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
Bedside Assessment	Burch, Seggie, and Gary (2006)	ZA/US	Medical	1 per week	Structured Feedback Form	Yes	Clinician Educator Facilitator	Rating, not contribute to grade
Case	Deane, Joyce, and Murphy (2015)	IE	Medical	5	Verbal feedback and feedback form	Yes		Rating, not contribute to grade
	Maley et al. (2008)	AU	Medical	No info	Model answers	No	No info	No info
Clickers	Palmer and Devitt (2014) (1)	AU	Medical	Dec-38	Model answers, explanations	No	No info	None
	Palmer and Devitt (2008)	AU	Medical	12	Detailed	No	No info	No info
	Cruz, Dias, and Kortemeyer (2011) (1)	BR	Physics	No info	No info	No	No info	None
	Gauci et al. (2009)	AU	Medical	No info	Correct answers and discussion	No	No info	None
Clinical Logbook	Hancock (2010)	US	Business	No info	No info	No	Teacher	Course credit
	Knight and Wood (2005)	US	Biology	Each lecture	Group discussion	No info	No info	Points, contribute to grade
	Nicol (2009) (2)	UK	French	No info	Group discussion	No info	No info	No info
Computer Aided Assessment	Paschal (2002) (1)	US	Engineering	No info	Discussion	No info	Teacher	Course credit
	Anziani, Durham, and Moore (2008)	UK	Dentistry	17	No info	Yes	Supervisor	Formative, not contribute to grade
Conferring Assessment	Gill and Greenhow (2008)	UK	Mechanics	No info	Elaborate, worked examples	No	No info	No info
	Baleghizadeh and Zarghami (2012)	IR	English	8	Content	Yes	Instructor	Formative, not contribute to grade
Data interpretation Exercises	English et al. (2006)	UK	Medical	1	Model answers	Yes	Peers/Teacher	~20% of grade
	Cacchione (2015)	ES	Italian	Two moments during class	Content comments and correction	No	Peers/Teacher	No info
	Buchanan (2000) (1)	UK	Psychology	3, 2 revision	No info	No info	No info	No info
	Buchanan (2000) (2)	UK	Psychology	5	No info	Yes	No info	No info
Formulate questions	Carnegie (2015)	CA	Medical	4	Content, worked example	No	No info	Points, contribute to grade
	Kennelly, Considine, and Flannery (2011)	IE	Economics	10	No info	No info	Teacher	25% of grade
	Krontiris-Litowitz (2008) (2)	US	Medical	1	No info	Yes	No info	No info
	Snowball and Mostert (2013) (3)	ZA	Economics	12	No info	No	No info	None
Formulate questions	Hardy et al. (2014)	UK	Science	At least 1	No	Yes	Teacher	2-5% of grade
	Zhang and Henderson (2015)	US	Medical	8	No	Yes	No info	None

(Continued)

Table 2. (Continued).

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
Game based assessment	Hooshyar et al. (2016) (1)	MY/TW/KR	Computer Science	1	Immediate knowledge of results	Yes	No info	Score, not contribute to grade
	Hooshyar et al. (2016) (2)	MY/TW/KR	Computer Science	1	Delayed knowledge of results	Yes	No info	Score, not contribute to grade
	Hooshyar et al. (2016) (3)	MY/TW/KR	Computer Science	1	Immediate elaborate	Yes	No info	Score, not contribute to grade
Homework	Schlegel and Selfridge (2014)	DM	Medical	No info	Formative	No info	Teacher	Points, no info about grade
	Cruz, Dias, and Kortemeyer (2011) (1)	BR	Physics	No info	No info	No	No info	None
	Cruz, Dias, and Kortemeyer (2011) (2)	BR	Physics	No info	No info	No	No info	2 bonus points on examination
Instructional Design Task	Kim and Ryu (2013)	KR	Teacher Education	1	Rubric and feedback messages	No info	Peers/Self	Rubric
	Hernick (2015)	US	Pharmacology	6	Content	No	No info	Formative, not contribute to grade
Midterm/examination/test/assessment	Azzi et al. (2015) (2)	CA	Medical	2	Immediate feedback	Yes	No info	10% of grade
	Brown et al. (2015) (2)	ZA/US	Medical	13/1	No info	Yes	Teacher	~21.7% of grade/2 pts per q
	Carrillo-de-la-Pena et al. (2009)	ES	Medical/Psychology/Biology	1	Correct answers	No	Teacher	Up to 5% bonus on grade
	Carrillo-de-la-Pena and Perez (2012)	ES	Psychology	3	Feedback on performance	No	No info	30% of grade
	Ernst and Colthorpe (2008) (1, 2)	AU	Medical	No info	Model answers and tutorial discussion	No	No info	Formative, not contribute to grade
	Fakcharoenphol, Potter, and Stelzer (2011)	US	Physics	No info	Correct answer or full solution	No	No info	No info
Learning Modules	Harwood (2006)	UK	No info	1	No info	No	No info	No info
	Henly (2003)	AU	Dentistry	After each educational meeting	Correct answers, model answers	No info	No info	No info
	Henly and Reid (2001)	AU	Science	After each educational meeting	Correct answers, model answers	Yes	No info	No info
	Jain, Agrawal, and Biswas (2012)	IN	Biochemistry	4	Correct answers, frequent errors, tips	No info	Teacher	No info
Kerdijk et al. (2013)	NL	Medical	3	Correct answers	Yes	Teacher	100% of grade	

(Continued)



Table 2. (Continued).

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
Midterm/ examination/test/ assessment	Krasne et al. (2006)	US	Medical	8	Correct/incorrect, explanations	7 required	No info	Score, not contrib- ute to grade
	Lawrence (2013) (1)	US	Psychology	3	No info	Yes	No info	No info
	Lawrence (2013) (2)	US	Psychology	1	No info	No	No info	Extra credit
	MacDermott (2013)	US	Economics	3	None	Yes	Teacher	60% of grade
	Micklewright et al. (2010)	US	Sport Science	7	Specific and generic class based	Yes	Teacher	100% of grade
	Nelson et al. (2009)	US	Biology	Weekly	Peer critique, teacher modelling	No info	No info	Points, not contrib- ute to grade
	Nicol (2009) (2)	UK	French	Fortnightly/1	No info	Some	Self and Teacher	~30% of grade
	Olson and McDonald (2004)	US	Dentistry	1	Correct answers	No	No info	Score, not contrib- ute to grade
	Palmer and Devitt (2014) (2)	AU	Medical	Weekly	None	Yes	No info	10% of grade
	Perdigones et al. (2009) (2)	ES	Engineering	Weekly (8)/ monthly	Correct answers	Yes	Teacher	Bonus points on final examination
	Ramirez (2010)	CL	Medical	1-Mar	No info	Yes/No	Teacher/Self	Graded/not con- tribute to grade
	Raupach et al. (2013) (1)	DE	Medical	1	No info	Yes	Teacher	1% of credit of the year
	Raupach et al. (2013) (2)	DE	Medical	1	No info	Yes	Teacher	Score, not contrib- ute to grade
	Rezaei (2015) (1,2,3)	US	Research Meth- ods	No info	No info	Yes	Teacher	Contributes to grade
	Sample et al. (2014)	US	Veterinary Med- icine	Weekly (6)	Model answers	Yes	Teacher	Scale, no info about grade
	Sargent and Curcio (2012) (2)	US	Law	1	No info	Yes	Teacher	Graded
	Siweya and Letsoalo (2014)	ZA	Chemistry	1/3-4/1	No info	Yes	Teacher	100% of grade (50/40/10)
	Snowball and Mostert (2013) (2)	ZA	Economics	2	No info	No	No info	30%
	Sullivan et al. (2013)	UK	Medical	1	None	No	Four examiners	No info
	Taradi, Taradi, and Radic (2005)	HR	Medical	10	Performance feed- back	No	Self	Not contribute to grade
	Tuunila and Pulkkinen (2015) (1)	FI	Chemical Engi- neering	5-Apr	General feedback, suitable answers	No	Teacher	Contributes to grade
	Tuunila and Pulkkinen (2015) (2)	FI	Chemical Engi- neering	6-Jan	General feedback, suitable answers	No	Teacher	Contributes to grade
	van der Kleij et al. (2012) (1)	NL	Economics	1	Immediate, elaborate	No	Computer	Score, not contrib- ute to grade
	van der Kleij et al. (2012) (2)	NL	Economics	1	Delayed, elaborate	No	Computer	Score, not contrib- ute to grade
	van der Kleij et al. (2012) (3)	NL	Economics	1	Delayed knowledge of results	No	Computer	Score, not contrib- ute to grade

(Continued)

Table 2. (Continued).

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
No info	Velan et al. (2008)	AU	Biomedical Science	1	No info	No	Self	No info
Oral Examination	Wiminger (2005) Perdigones et al. (2009) (1) Priya et al. (2012) (2)	US ES IN	Psychology Engineering Dentistry	1 No info 6	Class discussion No info No info	Yes No info Yes	Teacher No info Teacher	Extra credit No info Score, not contribute to grade
Portfolio	Behrouzi and Mowlavian (2015)	IR	English	12 entries	Conference with instructor	Yes	Instructor/Self	No info
Practice Attachment	Thome, Hovenberg, and Edgren (2006) Haffing, Beckman, and Edgren (2011)	SE SE	Biomedical Science Medical	3 2	Portfolio discussions Feedback form and narrative comments	Yes Yes	Teacher Tutor/Self	Scale, not contribute to grade Score, not contribute to grade
Pre-test/post-test	Lemus-Zuniga et al. (2015)	ES	Computer engineering	No info	Yes	No info	Teacher	No info
Presentation Project	Munson and Pierce (2015) Perdigones et al. (2009) (3) Rezaei (2015) (1,2,3)	US ES US	Pharmacology Engineering Research Methods	6 1 No info	Correct answers No info No info	No Yes Yes	No info Teacher/Peers Teacher	Score, not contribute to grade ~20% of grade No info
Quiz	Azzi et al. (2015) (1)	CA	Medical	17	Immediate feedback	Yes	No info	Formative, not contribute to grade
	Bijol, Byrne-Dugan, and Hoenig (2015)	US	Medical	5	Correct answers and explanations	No	Self	No info
	Bookhan, Warren, and Schoeman (2014)	ZA	Dentistry	2	Correct answers and group discussion	No	No info	% score, not contribute to grade
	Brown et al. (2015) (1) Cukusic, Garaca, and Jadric (2014)	ZA/US HR	Medical Economics	13 8	No info Correct/incorrect	Yes ≥60% on 6	Teacher Self	~11.6% of grade Score, not contribute to grade
	Dobson (2008)	US	Exercise Physiology	10	Yes	Yes	No info	10% of grade
	Domenech et al. (2015) Fauth (2015)	ES US	Economics Chemistry	10 Weekly	Suggested answers No info	Yes No info	Teacher No info	85% of grade No info
	Glass (2009) Ibabe and Jauregizar (2010)	US ES	Psychology Psychology	28 No info	Correct answers Correct/incorrect, hints	No info No	No info No info	Bonus points Score, not contribute to grade
	Jacoby et al. (2014)	UK	Biology	Weekly	Yes	No info	Self	Score, not contribute to grade

(Continued)



Table 2. (Continued).

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
Quiz	Kibble (2007) (1)	GD	Medical	2	No info	No	No info	Score, not contribute to grade
	Kibble (2007) (2)	GD	Medical	2	No info	No	No info	1% of grade, on login
	Kibble (2007) (3)	GD	Medical	2	No info	No	No info	2% of grade, score >30%
	Kibble (2007) (4)	GD	Medical	2	Incorrect answers shown	No	No info	2% of grade, based on actual score
	Kibble (2007) (5)	GD	Medical	2	Incorrect answers shown	No	No info	4% of grade, based on actual score
	Kibble et al. (2011)	US	Medical	Weekly (12)	Correct answers, faculty available for questions	No	No info	None
	Krontiris-Litowitz (2008) (1)	US	Neuroanatomy	2	No info	Yes	No info	No info
	Lawrence (2013) (3)	US	Psychology	After each chapter	No info	No	No info	No info
	Marden et al. (2013) (1,2)	AU	Medical	3	Individualised	Yes	Teacher	5% of grade
	Marden et al. (2013) (3)	AU	Medical	2	Individualised	Yes	Teacher	7.5% of grade
	Marden et al. (2013) (4)	AU	Medical	5	Individualised	Yes	Teacher	2% of grade, score >90%
	McNulty et al. (2015) (1)	US	Medical	3	Review session in class	No	Self	No info
	McNulty et al. (2015) (2)	US	Medical	After each lecture (14)	Instructive online feedback	No	Self	No info
	Mitra and Barua (2015)	MY	Medical	3	Content feedback	No info	No info	% score, not contribute to grade
	Paschal (2002) (2)	US	Biomedical Engineering	4	Correct answers	No info	Teacher	Course credit
	Rezaei (2015) (2-3)	US	Research Methods	No info	No info	Yes	Teacher	30% of grade
	Sargent and Curcio (2012) (1)	US	Law	5	Model answers and rubric	Yes	Teacher	Ungraded
	Williams, Bland, and Christie (2008) (1)	UK	Chemistry	20	Online feedback	No	Self	Formative, not contribute to grade
	Williams, Bland, and Christie (2008) (2)	UK	Chemistry	20	No info	No	Teacher	5% of grade, for total of quizzes
	Wilson et al. (2011)	CA	Geography	After each class	Correct/incorrect	No	Self	None

(Continued)

Table 2. (Continued).

Assessment Type	Article	Country	Discipline	Frequency	Feedback	Mandatory	Assessor	Scoring
Simulated Patient	Zhang and Henderson (2015) (1,2) Gums, Kleppinger, and Urick (2014)	US US	Medical Pharmacology	10 1	Correct answers, open feedback Rubric and talking points for indi- vidual	Yes Yes	Teacher Facilitator	None Rubric, not contrib- ute to grade
Writing Assignment	Birjandi and Hadidi Tamjid (2012) (1)	IR	Teaching English	Each session (12)	Oral or written teacher feedback, self-reflection Rubric	Yes	Teacher/Self	No info
	Birjandi and Hadidi Tamjid (2012) (2)	IR	Teaching English	No info	Rubric	Yes	Self/Teacher	Rubric, not contrib- ute to grade
	Birjandi and Hadidi Tamjid (2012) (3)	IR	Teaching English	No info	Rubric	Yes	Peers/Teacher	Rubric, not contrib- ute to grade
	Birjandi and Hadidi Tamjid (2012) (4)	IR	Teaching English	No info	Rubric	Yes	Self/Peers	Rubric, not contrib- ute to grade
	Birjandi and Hadidi Tamjid (2012) (5)	IR	Teaching English	No info	No info	Yes	Teacher	No info
	Covic and Jones (2008)	AU	Psychology	1 (resubmission)	Comments and suggestions, Form, Group feedback	No	Teacher	Score, 20% of grade
	Fisher, Cavanagh, and Bowles (2011)	AU	Management	1 (draft)	Feedback sheet and discussion	No	Teacher	None
	Gielen and De Wever (2015)	BE	Educational Science	3 (drafts)	Feedback form	Yes	Peers	Rubric, not contrib- ute to grade
	Isaksson (2008)	SE	Archaeology	10	Short written feed- back	Yes	Lecturer	100% of grade
	Mulder et al. (2014)	AU	Animal Behaviour	1 (draft)	Structured review sheet	Yes	Peers	Score, not contrib- ute to grade
	Nicol (2009) (1)	UK	Psychology	6 light, 6 deep	Peer discussion, model answers afterwards	Yes	Peers/Self	No info
	Perdigones et al. (2009) (3)	ES	Engineering	1	No info	Yes	Journal/Teacher	30% of grade
	Priya et al. (2012) (1)	IN	Dentistry	6	No info	Yes	Teacher	Score, not contrib- ute to grade
	Snowball and Mostert (2013) (1)	ZA	Economics	2	Rubric	No	Peers/Teacher	30% of grade
	Stellmack et al. (2012)	US	Psychology	1 (resubmission)	Rubric	No info	Peers/Self/ Teacher	Half points for draft
	Wakefield et al. (2014)	UK	Sport Studies	1	Feedback checklist or marking sheet	Yes	Teacher	60% of grade

Note: Bolded references have found a positive relation between the assessment and student grades.

Generally, research on intermediate assessment falls into two categories. The first category is correlational studies addressing the relationship of the intermediate assessment and final assessment grades ($N = 16$). The second category includes studies that compare final examination scores for groups with and without intermediate assessment, or scores on different courses ($N = 62$). Furthermore, fourteen studies investigated a correlation as well as a comparison. In the 88 selected articles, 77 studies found positive results of intermediate assessment, and 23 found no results. To reiterate, several articles discussed multiple outcomes, which is why the total number of results exceeds 88.

Assessment type

In total, 24 different types of assessment are used in the 88 articles. Table 2 shows that the most prevalent assessment types are midterms/examinations, quizzes and writing assignments. These three are all well-known, 'generic', conventional assessment forms, but the range of assessments also includes game-based assessments, clinical bedside assessments, and the use of twitter as a dynamic assessment tool.

Assessment type and discipline

When disciplines are aggregated to a more general level it becomes evident that midterms (or related generic forms of examinations) and writing assignments are the only types of assessment used across all disciplines; Quizzes are also represented across the majority of disciplines. These three generic assessment types can be implemented in all disciplines. There are also some discipline specific assessment types, like clinical bedside assessment for medical students. Interestingly, the oral presentation, which can be considered a generic assessment type as well, was only used in one study in one discipline.

Assessment type and frequency of assessment

Students are confronted with a vastly differing amount of assessments, as can be seen in the frequency column of Table 2. The majority of authors describe courses that consist of one to ten intermediate assessments per course. Palmer and Devitt (2014) mention the highest number of intermediate assessments in one course: students had access to 38 intermediate assessments.

Since not all courses run for the same amount of time, and eight assessments in an eight week course will be perceived differently than eight assessments in a 20 week course, direct comparison of assessment frequency in different studies may be biased. Some authors specify that their assessments occur weekly (e.g. Nelson et al. 2009), fortnightly (e.g. Nicol 2009) or after each educational meeting (e.g. Henly and Reid 2001).

Overall, it can be argued that the frequency of assessments is not the most important characteristic, but rather the alignment between the frequency of assessment and the course goals. For example, the 38 assessments in Palmer and Devitt (2014) may seem like a lot, but they reported completion rates of 90%, indicating that the number of assessments was suitable for the course.

Assessment type and feedback

The feedback column in Table 2 indicates that the majority of authors report some form of feedback. The elaborateness of the feedback ranges from simply stating whether an answer was correct, to giving elaborate, content-related, qualitative feedback. It seems that some types of feedback are closely related to specific types of assessment, which is most clearly the case for the rubric or feedback form, which is mainly employed in more open tasks like writing assignments and patient encounters. Interestingly, it seems that the category of feedback which one expects to have the largest influence (elaborated content feedback) does not produce more positive effects than supplying students with their results or the correct answers. This may have to do with different feedback effects for different student groups. For example, Shute (2008) poses that while elaborated feedback is the optimal feedback type for lower

achieving students, high achieving students often get enough information from corrective feedback. This is also in line with work by Brookhart (2001), who discovered that successful students often use summative test information for learning purposes.

Two papers that illustrate how the level of feedback can relate to student results are Hooshyar et al. (2016) and van der Kleij et al. (2012). Both of these papers compare conditions with elaborate feedback to feedback that only consist of knowledge of results.

Hooshyar et al. (2016) investigated an online game based formative assessment (a tic-tac-toe variation, where a correct answer on a question allows the student to place a token on the grid) and its influence on problem solving skills. The population was 52 first year computer science students, 22 in a control class working with what the authors call 'traditional approaches' (Hooshyar et al. 2016, 27), and 30 in the experimental class working with the game-based assessment. There were three feedback conditions in the game-based assessment: delayed knowledge of results, immediate knowledge of results, and immediate elaborate feedback. Results indicate that students in the experimental condition performed better than their peers in the control group, and that immediate elaborate feedback proved to be the most useful.

The research by van der Kleij et al. (2012) focused on a computer-based assessment for learning task for 152 first year economics students. The study compared summative assessment results of three groups. Group one consisted of 52 students who received immediate knowledge of correct response and elaborate feedback on their performance. Group two ($N = 48$) received delayed knowledge of correct response and elaborate feedback, and group three ($N = 52$) only received delayed knowledge of their results. This intervention had no effect on the final summative assessment results, indicating immediate or delayed elaborate feedback did not improve final performance.

The studies by Hooshyar et al. (2016) and van der Kleij et al. (2012) exhibit a different influence of feedback, where the former does find an influence of elaborate feedback and the latter does not. There are several similarities in the design of the two studies. Both focus on a single assessment for first year students, both use multiple-choice questions and both assessments are performed on a computer. The main difference is that for Hooshyar et al. the assessment was a mandatory part of the course, following a specific learning module on problem solving. In the case of van der Kleij et al., the assessment for learning was mainly part of the experiment, and participation in the assessment was not a prerequisite of participating in the course.

Assessment type and mandatory assessments

There is an even divide in the use of mandatory versus voluntary intermediate assessment. Interestingly, whether or not an assessment is mandatory does not immediately seem to influence students' results. For mandatory and voluntary assessments, the amount of studies finding no positive influence of intermediate assessment is approximately 36% and 31% respectively.

Carrillo-de-la-Peña and Perez (2012) used mandatory intermediate assessment in an interesting way. Their research focused on three classes of second year psychology students for three consecutive years, with a total population of 903 students. These students were given the choice to partake in continuous assessment (CA) in their course, or to stick to 'traditional' assessment (TA). Although the choice for CA was voluntary, the 30% of students who chose CA were then required to participate in mandatory formative assessments and 'complementary individual and group tasks' (46). At the end of the term all students took the same final examination, which counted for 70% for CA students and 100% for TA students.

Comparison of the groups shows that fewer CA students dropped the course during the semester, more CA students passed the examination, and CA students got higher marks for the examination. These results were consistently found for all three academic years under investigation. Surprisingly, TA students remarked in a questionnaire that they believed the CA system would have helped their learning and 33 to 60% of them would have chosen it, if they were starting the course now.

Snowball and Mostert (2013) described voluntary participation in intermediate assessment for a writing assignment. During their course, Snowball and Mostert (2013) were faced with technical difficulties

during the period of online submission of the essay assignment for peer assessment, which led them to also offer an alternative exercise, making the peer assessment voluntary. This alternative led to only about half of the 800 students enrolled in the macroeconomics class participating in peer assessment. Regression analyses showed that participating in the peer assessment did not predict students' final essay grades. Even though 58% of the students agreed that peer feedback helped them to improve the quality of their essay, only 10% of the students made significant changes to their essays. Furthermore, when comparing peer assessment on the first version with tutor assessment of the final version, it seemed that peers were marking in a much smaller range (between 60 and 68%) than the tutors were, which may not have motivated students to implement significant changes.

In summary, research does not show differences between mandatory and voluntary assessments with respect to students' grades. However, in the case of voluntary assessments, selection effects should be taken into account. For example, Covic and Jones (2008) gave third year psychology students the opportunity to resubmit their 1200 word essays; however, high performing students chose not to participate in the resubmission. In other cases it may be that high performers do participate, whereas lower performing students do not. This begs the question of what other variables influence the role of intermediate assessment. One that clearly comes forward from Snowball and Mostert's research is 'who is the assessor?', and another is 'what can students gain with participating in the assessment?'

Assessment type and assessor

The majority of authors describe assessments that were assessed by the teacher (or a lecturer, facilitator, supervisor, etcetera), or did not explicitly name who the assessor was. Research that uses peer-assessment or peer feedback is mainly limited to writing assignments, whereas the majority of authors that employ self-assessment do so using quizzes.

One study that compares the differences between self, peer, and teacher-assessment is Birjandi and Hadidi Tamjid (2012). In this paper, the researchers investigate the writing performance of 157 Teaching English as a Foreign Language students, divided into four experimental groups and one control group. Students in all five groups wrote essays and subsequently were assessed on those essays in different ways. In the first group, teachers gave oral feedback when students read their written essays aloud in class, and students wrote reflection journals detailing what they learned and what problems they had during the class. In the second group, students self-assessed their essays using a rubric and were able to make revisions based on this assessment. Afterwards, teachers assessed the essays and students were able to compare their self-assessment to that of the teacher. The third group was set up very similarly to the second group, but instead of self-assessment, this group was divided into smaller peer groups where students rated each other's work using the rubric. The fourth group was the only group that did not use teacher assessment. In this group students first assessed themselves and subsequently went into peer groups to assess the work of their peers. The fifth group was the control group, where teachers assessed the essays.

Before and after writing their essays, students took an English proficiency test and a writing test, to see if their performance improved. For all five groups, scores on the writing test increased after the classes on essay writing. Analysis of variance with post-hoc tests indicated that the third group (peer and teacher assessment) scored significantly better than the first, second and fifth groups.

These findings, together with several other articles using peer assessment, suggest a positive influence of peer assessment. However, the influence of teacher assessment should not be excluded. When comparing the different groups in Birjandi and Hadidi Tamjid's (2012) research, the fourth group (using self and peer assessment) did not differ from the other groups, indicating that the results of the third group are not just due to peer assessment, but to the combination of teacher and peer assessment. When combining two or more assessors, the teacher is often the assessor who decides on the final evaluation of the assessment: in most assessments the teacher is still the decider.

Assessment type and reward

Another important factor in assessment is the type of reward students get for participating. In the scoring column of Table 2 the two largest categories contain assessments that either make up a percentage of the final grade, or assessments on which students simply receive a (formative) score, that does not count towards the final grade. In some cases, teachers gave students bonus points for completing the assessments. Out of the two biggest categories the most positive results are found in studies awarding a percentage of course grade, maybe indicating that students who feel that they have something to gain with the assessment will exert more effort.

Kibble (2007) provides an interesting example of how rewards seem to influence the effectiveness of an intermediate assessment. Five cohorts of medical physiology students could participate in two voluntary quizzes as a means of formative assessment. In the first cohort, students got no rewards for participating in the quizzes. In the second cohort, students could earn 0.5% course credit per quiz simply for taking the quiz, irrespective of their grade. Students in cohort three received 1% credit per quiz if they scored >30%, and they were free to retake the quiz as often as they wanted in a one-week period. In cohorts four and five students were rewarded with 1 or 2 percent course credit, 'allotted as a function of the actual quiz score attained from the better of two attempts' (254) respectively.

As Table 2 indicates, in cohorts four and five there was no longer a positive effect of participating in the quizzes. However, participation in the quizzes increased drastically, from 52% in the first to 98% in the fifth cohort, when the rewards were greater. In all cohorts the summative assessment scores of quiz takers were compared to non-takers, which may explain the lack of difference in cohorts four and five, since the group of non-takers was small. Furthermore, results seem to indicate that students in cohorts four and five did not engage with the quizzes with the formative purpose in mind, but mainly to gain course credit. Kibble's main evidence for this is that these students got a very high score on their first attempt at the quiz, subsequently did not try to retake the quiz for improvement, and were unable to sustain their high performance in the summative examination.

This may be evidence to be cautious with rewards for intermediate assessments with a formative purpose. Gibbs and Simpson (2004, 23) mention that assessment only supports learning when the provided feedback is 'received and attended to', subsequently showing several research outcomes that indicate students often discard feedback and only attend to grades when they are provided together. However, grades or bonus points can work as an incentive for students to participate in the intermediate assessment. Harland et al. (2015) stated that students are focused on working for graded assessments more than for ungraded educational activities.

Intermediate assessment scenarios

The review shows how various characteristics of intermediate assessments are related to student grades. In this section, three scenarios for intermediate assessment are presented, which indicate combinations of the characteristics discussed in the results section. These scenarios illustrate how the different characteristics discussed could be implemented in a higher education setting. Each scenario will focus on a different assessment type: the quiz, the written assignment and the authentic professional assignment. These three types reflect different aspects of the intermediate assessment process. In the case of all these scenarios it is important to keep in mind that different assessment goals may ask for different assessment characteristics.

Scenario 1: quiz

Quizzes are a frequently used assessment type. In the current scenario, the quiz is implemented in a first-year social science course. This course has weekly lectures and lasts for ten weeks. Teachers felt that students who are attending the lectures often come unprepared, so they decided to implement quizzes to keep the students on track in their learning. These quizzes are not mandatory, because the

teachers also want to facilitate students' ownership over their learning, but participation in the quizzes is rewarded with a 0.5 (out of 10) bonus when students pass all quizzes. The student population of this course is fairly large, with about 400 students enrolled.

Quizzes take place after each lecture, to encourage students to keep up with their studying. A quiz after each lecture clearly aligns the assessment with the materials studied during the lecture. For students taking a 10–20 item quiz, consisting mainly of multiple choice questions and some short answer questions, is also a good instrument for self-assessment and to check whether they understand all the class materials.

To alleviate teacher workload, the quizzes are implemented using an online platform, with the additional bonus that students can access the quizzes at their own convenience. In the online quiz, the assessment process is automated and students are supplied with feedback. This feedback is tailored to the student's level of performance, since Shute (2008) suggests that low achieving students benefit from elaborated feedback, whereas for high achieving students simple corrective feedback is often enough. Elaborated feedback on the multiple choice questions consists of explanations on each answer option, including why the correct choice is correct. For the short answers questions, a model answer with the most important parts of the correct answer will be supplied after answering the question. Because the computer cannot adequately assess the correctness of short answer questions, only the multiple choice questions are used when deciding whether or not a student passed the quiz.

Reflections on the quiz scenario

Several advantages of the quiz are: they are often easy to design and use, administering them is time-efficient, they are fit for use in large groups, they help keep students on track, they promote active learning, and can reduce test anxiety. In the scenario the main two advantages were the ease of large group assessment and the ability to keep students on track.

An advantage of hosting the quizzes on an online platform is that questions for each week can be randomly chosen from a larger bank of quiz items. Quizzes that are composed by selecting items from a larger item bank allow meaningful repetition of the quiz more easily, indicating that students learn the associated learning materials instead of simply memorising the correct answers on a quiz. However, building a large item bank is a labour intensive process. Furthermore, if the item bank is large enough, there may be no new questions available for the final examination.

One disadvantage of intermediate quizzes is strategic quiz taking, as was indicated when discussing the research by Kibble (2007). Simply memorising quiz answers, or other surface learning approaches for answering multiple choice questions, often does not promote long term learning effects (Gibbs and Simpson 2004). This can be overcome by integrating a discussion of quiz answers in the following lecture.

We believe that the majority of the characteristics described in Scenario 1 can be applied to scenarios containing a midterm or formative examination as intermediate assessment as well. The main difference would be that examinations usually occur less often, are larger in size than quizzes are, and usually have higher stakes than quizzes do. When midterms are used to complete certain topics that will not return on the final examination, it may be of less importance to offer elaborate feedback.

Scenario 2: writing assignment

Writing assignments were also among the most often used assessment types. This scenario concerns a twelve week third-year business course of 80 students, where one of the final assessments is a 3000-word business plan. The intermediate assessment that is part of this scenario is a peer assessment. Students are invited to voluntarily submit a draft of their business plan as an opportunity to receive formative peer feedback.

In the second week of the course, before students start writing their plans, they are invited to participate in a peer feedback workshop. During this workshop the teacher introduces them to the procedure of peer feedback, using an online peer feedback platform. Furthermore, during this workshop, the students familiarise themselves with the assessment criteria. All students are asked to provide feedback

on a business plan written by a student in a previous cohort, using a proposed feedback form. After this exercise the teacher and students discuss the criteria and amend them where necessary.

After the peer feedback workshop, students have two weeks to finish their draft for peer feedback. Following submission of the draft, the online platform automatically matches two students to give feedback on each other's drafts. The teacher decided against predetermined couples matched on students' ability, since the voluntary nature of the draft submission meant that there was a risk that several couples would remain incomplete. Students have two weeks to complete the peer feedback assignment. When students have received their drafts with feedback, they subsequently have four weeks to finish their business plan and submit the final version to their teachers.

Reflections on the writing assignment scenario

This type of assessment is versatile and is often used to assess writing skill and content knowledge or discipline specific skills simultaneously. Writing assignments often allow students to display their in-depth knowledge and present a good opportunity to implement peer assessment in the curriculum. Peer assessment is a valuable learning tool, where students learn from the feedback they receive, as well as from providing feedback to their peer. Furthermore, adopting peer feedback in the writing assignment provides the students with an opportunity to receive feedback on their work, without significantly increasing teacher workload.

It is instrumental that students receive adequate peer feedback training (van Zundert, Sluijsmans, and van Merriënboer 2010). Van Zundert et al. found that peer feedback training positively influenced the quality of peer feedback as well as students' attitudes towards peer assessment. Furthermore, during this training students can familiarise themselves with the assessment criteria and propose to amend them where they feel that they do not fit the assignment.

Results from Snowball and Mostert (2013) indicate two possible problems with peer assessment on writing. First, students did not make significant improvements based on the peer feedback. Second, peers marked on a much smaller range than teachers did, indicating they were less critical, but also less positive. The latter problem could be overcome by training and experience.

Scenario 3: authentic assessment

Even though authentic assessments were not prevalent in the current sample, they do make up an important facet of the assessment domain. In courses in, for example, the medical field more conventional written assessments are often unable to assess the professional skills that are needed in the workplace, necessitating assessment types that simulate these practical situations, or sometimes even take place in practice.

This scenario concerns an intermediate assessment in a semester long fourth year medical course, attended by 35 students. This course focuses on clinical skills development and the final examination in this course is a 10 station objective structured clinical examination (OSCE). Since clinical skills cannot be taught by purely lecturing, this course features a weekly practical tutorial in addition to the lectures focusing on theory. These tutorials are small scale, with each tutorial only consisting of seven students. To fully prepare students for the examination, all stations that are included in the OSCE are administered formatively during the tutorials, at a rate of one station every other week.

The formative OSCE focuses on the stations that were discussed during lectures and tutorials that week. All students perform the tasks associated with the station on a simulated patient, while being observed by their peers and a tutor. The tutor subsequently judges students' performance on the station by completing a rating form. On this form there is ample space for the tutor to provide narrative feedback on the students' performance. Peers have no formal assessment role, but they are invited to share their observations and tips.

Reflections on the authentic assessment scenario

Authentic assessments help students experience real life (or close to real life) situations where they need to apply their skills. However, these assessments are often time consuming.

An important facet of the current scenario is that students are learning from each observing each other. This does mean that the first student who completes the station may be at a disadvantage, because they have not observed a peer. It is up to the tutor to ensure that different students are the first to participate in the station each assessment. Furthermore, teachers can model the correct performance or show video examples of experts.

Concluding remarks

This review set out to identify which characteristics of intermediate assessment may play a role in improving students' final examination results. The results of the review show that intermediate assessment is employed widely in higher education, and that the variety is large. The instructional designs vary per study, depending on the educational context and the specific objectives the instructors had in mind.

The main outcomes indicate the following:

- (1) Corrective feedback is often enough, but lower achieving students may need more elaborate feedback.
- (2) Both mandatory and voluntary assessments can influence students' grades. Therefore, intermediate assessment is not only effective when it is mandatory, but can also be effective when it is voluntary.
- (3) Peer assessment is beneficial to students' grades, especially since providing feedback is a learning opportunity for students as well.
- (4) Rewarding students increases their participation, but may negatively influence the learning effects of the assessment when students are strictly focused on the reward and do not engage with feedback.

In a way, intermediate assessment characteristics are different pieces of a puzzle that, combined together, influence students' grades. Teachers can optimise their courses by checking which characteristics align with their goals, and combining these characteristics in such a way that students can achieve the educational learning goals more easily and perform better.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Indira N. Z. Day is a PhD candidate at the Leiden University Graduate School of Teaching. Her research focuses on characteristics of intermediate assessment and study outcomes in higher education.

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

Michiel Westenberg is full professor of developmental psychology. 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 full professor of educational sciences and the 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>

Wilfried Admiraal  <http://orcid.org/0000-0002-1627-3420>

References

- Articles marked by an * were included in the literature review.
- 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.
- *Anziani, H., J. Durham, and U. Moore. 2008. "The Relationship between Formative and Summative Assessment of Undergraduates in Oral Surgery." *European Journal of Dental Education* 12 (4): 233–238. doi:10.1111/j.1600-0579.2008.00524.x.
- *Azzi, A. J., C. J. Ramnanan, J. Smith, E. Dionne, and A. Jalali. 2015. "To Quiz or Not to Quiz: Formative Tests Help Detect Students at Risk of failing the Clinical Anatomy Course." *Anatomical Sciences Education* 8 (5): 413–420. doi:10.1002/ase.1488.
- *Baleghizadeh, S., and Z. Zarghami. 2012. "Impacto de la evaluación mediante conferencias en el aprendizaje de la gramática en estudiantes de inglés como lengua extranjera [The Impact of Conferencing Assessment on EFL Students' Grammar Learning]." *Profile Issues in Teachers' Professional Development* 14 (2): 131–144.
- Bälter, O., E. Enström, and B. Klingenberg. 2013. "The Effect of Short Formative Diagnostic Web Quizzes with Minimal Feedback." *Computers & Education* 60 (1): 234–242. doi:10.1016/j.compedu.2012.08.014.
- *Behrouzi, P., and R. Mowlavian. 2015. "The Effect of Assessment Portfolios on the Paragraph Writing Development of Iranian Intermediate EFL Learners." *Modern Journal of Language Teaching Methods* 5 (3): 306–320.
- 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.
- Biggs, J. 1996. "Enhancing Teaching through Constructive Alignment." *Higher Education* 32 (3): 347–364. doi:10.1007/BF00138871.
- *Bijol, V., C. J. Byrne-Dugan, and M. P. Hoenig. 2015. "Medical Student Web-based Formative Assessment Tool for Renal Pathology." *Medical Education Online* 20: 26765. doi:10.3402/meo.v20.26765.
- *Birjandi, P., and N. Hadidi Tamjid. 2012. "The Role of Self-, Peer and Teacher Assessment in Promoting Iranian EFL Learners' Writing Performance." *Assessment & Evaluation in Higher Education* 37 (5): 513–533. doi:10.1080/02602938.2010.549204.
- Black, P., and D. Wiliam. 2004. "The Formative Purpose: Assessment Must First Promote Learning." *Yearbook of the National Society for the Study of Education* 103 (2): 20–50. doi:10.1111/j.1744-7984.2004.tb00047.x.
- *Bookhan, V., N. Warren, and H. Schoeman. 2014. "Do Precursor Tests Influence the Performance of Oral Hygiene Students in Subsequent Semester Tests?" *SADJ: Journal of the South African Dental Association* 69 (10): 464–467.
- Brookhart, S. M. 2001. "Successful Students' Formative and Summative Uses of Assessment Information." *Assessment in Education: Principles, Policy & Practice* 8 (2): 153–169. doi:10.1080/09695940123775.
- *Brown, G. A., M. R. Bice, B. S. Shaw, and I. Shaw. 2015. "Online Quizzes Promote Inconsistent Improvements on in-class Test Performance in Introductory Anatomy and Physiology." *Advances in Physiology Education* 39 (2): 63–66. doi:10.1152/advan.00064.2014.
- *Buchanan, T. 2000. "The Efficacy of a World-wide Web Mediated Formative Assessment." *Journal of Computer Assisted Learning* 16 (3): 193–200. doi:10.1046/j.1365-2729.2000.00132.x.
- *Burch, V. C., J. L. Seggie, and N. E. Gary. 2006. "Formative Assessment Promotes Learning in Undergraduate Clinical Clerkships." *South African Medical Journal* 96 (5): 430–433.
- *Cacchione, A. 2015. "Creative Use of Twitter for Dynamic Assessment in Language Learning Classroom at the University." *Interaction Design and Architectures* 24: 145–161.
- *Carnegie, J. 2015. "Use of Feedback-oriented Online Exercises to Help Physiology Students Construct Well-organized Answers to Short-answer Questions." *CBE-Life Sciences Education* 14 (3): 1–12. doi:10.1187/cbe-14-08-0132.
- *Carrillo-de-la-Peña, M. T., and J. Perez. 2012. "Continuous Assessment Improved Academic Achievement and Satisfaction of Psychology Students in Spain." *Teaching of Psychology* 39 (1): 45–47. doi:10.1177/0098628311430312.
- *Carrillo-de-la-Peña, M. T., E. Bailles, X. Caseras, A. Martinez, G. Ortet, and J. Perez. 2009. "Formative Assessment and Academic Achievement in Pre-graduate Students of Health Sciences." *Advances in Health Sciences Education* 14 (1): 61–67. doi:10.1007/s10459-007-9086-y.
- Clouder, L., C. Broughan, S. Jewell, and G. Steventon. 2012. *Improving Student Engagement and Development through Assessment*. London: Routledge.
- *Covic, T., and M. K. Jones. 2008. "Is the Essay Resubmission Option a Formative or a Summative Assessment and Does It Matter as Long as the Grades Improve?" *Assessment & Evaluation in Higher Education* 33 (1): 75–85. doi:10.1080/02602930601122928.
- *Cruz, É., H. Dias, and G. Kortemeyer. 2011. "Efeito da avaliação formativa em cursos de física em universidades brasileiras [The Effect of Formative Assessment in Brazilian University Physics Courses.]" *Revista Brasileira de Ensino de Física* 33 (4): 4315–4315. doi:10.1590/s1806-11172011000400016.
- *Cukusic, M., Z. Garaca, and M. Jadric. 2014. "Online Self-assessment and Students' Success in Higher Education Institutions." *Computers & Education* 72: 100–109. doi:10.1016/j.compedu.2013.10.018.
- Day, I. N. Z., F. M. van Blankenstein, P. M. Westenberg & W. F. Admiraal. 2016. "Literatuurreview: Kenmerken van tussentijds toetsen in het hoger onderwijs en hun relatie met opbrengst bij studenten." [Literature review: Characteristics of

- intermediate assessment in higher education and their relation to student outcomes.] Roundtable Session Presented at the Annual Meeting of the Netherlands Educational Research Association, Rotterdam, the Netherlands, May 25–27.
- Day, I. N. Z., F. M. van Blankenstein, P. M. Westenberg, and W. F. Admiraal. 2017. "Teacher and Student Perceptions of Intermediate Assessment in Higher Education." *Educational Studies*. Advance online publication. doi:10.1080/03055698.2017.1382324
- De Paola, M., and V. Scoppa. 2011. "Frequency of Examinations and Student Achievement in a Randomized Experiment." *Economics of Education Review* 30 (6): 1416–1429. doi:10.1016/j.econedurev.2011.07.009.
- *Deane, R. P., P. Joyce, and D. J. Murphy. 2015. "Team Objective Structured Bedside Assessment (TOSBA) as Formative Assessment in Undergraduate Obstetrics and Gynaecology: A Cohort Study." *BMC Medical Education* 15 (172): 1–12. doi:10.1186/s12909-015-0456-5.
- *Dobson, J. L. 2008. "The Use of Formative Online Quizzes to Enhance Class Preparation and Scores on Summative Exams." *Advances in Physiology Education* 32 (4): 297–302. doi:10.1152/advan.90162.2008.
- *Domenech, J., D. Blazquez, E. de la Poza, and A. Munoz-Miquel. 2015. "Exploring the Impact of Cumulative Testing on Academic Performance of Undergraduate Students in Spain." *Educational Assessment Evaluation and Accountability* 27 (2): 153–169. doi:10.1007/s11092-014-9208-z.
- 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 (1): 4–58. doi:10.1177/1529100612453266.
- *English, R., S. T. Brookes, K. Avery, J. M. Blazeby, and Y. Ben-Shlomo. 2006. "The Effectiveness and Reliability of Peer-marking in First-year Medical Students." *Medical Education* 40 (10): 965–972. doi:10.1111/j.1365-2929.2006.02565.x.
- *Ernst, H., and K. Colthorpe. 2008. "Expanding Voluntary Active-learning Opportunities for Pharmacy Students in a Respiratory Physiology Module." *American Journal of Pharmaceutical Education* 72 (2):28.
- *Fakcharoenphol, W., E. Potter, and T. Stelzer. 2011. "What Students Learn When Studying Physics Practice Exam Problems." *Physical Review Special Topics-physics Education Research* 7 (1): 010107. doi:10.1103/PhysRevSTPER.7.010107.
- *Fautch, J. M. 2015. "The Flipped Classroom for Teaching Organic Chemistry in Small Classes: Is It Effective?" *Chemistry Education Research and Practice* 16 (1): 179–186. doi:10.1039/c4rp00230j.
- *Fisher, R., J. Cavanagh, and A. Bowles. 2011. "Assisting Transition to University: Using Assessment as a Formative Learning Tool." *Assessment & Evaluation in Higher Education* 36 (2): 225–237. doi:10.1080/02602930903308241.
- *Gauci, S. A., A. M. Dantas, D. A. Williams, and R. E. Kemm. 2009. "Promoting Student-centered Active Learning in Lectures with a Personal Response System." *Advances in Physiology Education* 33 (1): 60–71. doi:10.1152/advan.00109.2007.
- Gibbs, G., and C. Simpson. 2004. "Conditions under Which Assessment Supports Students' Learning." *Learning and Teaching in Higher Education* 1 (1): 3–31.
- *Gielen, M., and B. De Wever. 2015. "Structuring the Peer Assessment Process: A Multilevel Approach for the Impact on Product Improvement and Peer Feedback Quality." *Journal of Computer Assisted Learning* 31 (5): 435–449. doi:10.1111/jcal.12096.
- *Gill, M., and M. Greenhow. 2008. "How Effective is Feedback in Computer-aided Assessments?" *Learning Media and Technology* 33 (3): 207–220. doi:10.1080/17439880802324145.
- *Glass, A. L. 2009. "The Effect of Distributed Questioning with Varied Examples on Exam Performance on Inference Questions." *Educational Psychology* 29 (7): 831–848. doi:10.1080/01443410903310674.
- *Gums, T. H., E. L. Kleppinger, and B. Y. Urick. 2014. "Outcomes of Individualized Formative Assessments in a Pharmacy Skills Laboratory." *American Journal of Pharmaceutical Education* 78 (9): 166. doi:10.5688/ajpe789166.
- *Haffling, A.-C., A. Beckman, and G. Edgren. 2011. "Structured Feedback to Undergraduate Medical Students: 3 Years' Experience of an Assessment Tool." *Medical Teacher* 33 (7): E349–E357. doi:10.3109/0142159x.2011.577466.
- *Hancock, T. M. 2010. "Use of Audience Response Systems for Summative Assessment in Large Classes." *Australasian Journal of Educational Technology* 26 (2): 226–237.
- *Hardy, J., S. P. Bates, M. M. Casey, K. W. Galloway, R. K. Galloway, A. E. Kay, P. Kirsop, and H. A. McQueen. 2014. "Student-generated Content: Enhancing Learning through Sharing Multiple-choice Questions." *International Journal of Science Education* 36 (13): 2180–2194. doi:10.1080/09500693.2014.916831.
- 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 (4): 528–541. doi:10.1080/02602938.2014.931927.
- *Harwood, I. 2006. "Managing Summative Computer-aided Assessment Failures: Further Evidence for 'Optional Substitution.'" *British Journal of Educational Technology* 37 (4): 637–641. doi:10.1111/j.1467-8535.2006.00583.x.
- *Henly, D. C. 2003. "Use of Web-based Formative Assessment to Support Student Learning in a Metabolism/Nutrition Unit." *European Journal of Dental Education* 7 (3): 116–122. doi:10.1034/j.1600-0579.2003.00310.x.
- *Henly, D. C., and A. E. Reid. 2001. "Use of the Web to Provide Learning Support for a Large Metabolism and Nutrition Class." *Biochemistry and Molecular Biology Education* 29 (6): 229–233.
- *Hernick, M. 2015. "Test-enhanced Learning in an Immunology and Infectious Disease Medicinal Chemistry/Pharmacology Course." *American Journal of Pharmaceutical Education* 79 (7):97.
- Holmes, N. 2015. "Student Perceptions of Their Learning and Engagement in Response to the Use of a Continuous E-assessment in an Undergraduate Module." *Assessment & Evaluation in Higher Education* 40 (1): 1–14. doi:10.1080/02602938.2014.881978.

- *Hooshyar, D., R. B. Ahmad, M. Yousefi, M. Fathi, S. J. Horng, and H. Lim. 2016. "Applying an Online Game-based Formative Assessment in a Flowchart-based Intelligent Tutoring System for Improving Problem-solving Skills." *Computers & Education* 94: 18–36.
- *Ibabe, I., and J. Jauregizar. 2010. "Online Self-assessment with Feedback and Metacognitive Knowledge." *Higher Education* 59 (2): 243–258. doi:10.1007/s10734-009-9245-6.
- *Isaksson, S. 2008. "Assess as You Go: The Effect of Continuous Assessment on Student Learning during a Short Course in Archaeology." *Assessment & Evaluation in Higher Education* 33 (1): 1–7. doi:10.1080/02602930601122498.
- *Jacoby, J. C., S. Heugh, C. Bax, and C. Branford-White. 2014. "Enhancing Learning through Formative Assessment." *Innovations in Education and Teaching International* 51 (1): 72–83. doi:10.1080/14703297.2013.771970.
- *Jain, V., V. Agrawal, and S. Biswas. 2012. "Use of Formative Assessment as an Educational Tool." *Journal of Ayub Medical College, Abbottabad* 24 (3–4): 68–70.
- *Kennelly, B., J. Considine, and D. Flannery. 2011. "Online Assignments in Economics: A Test of Their Effectiveness." *Journal of Economic Education* 42 (2): 136–146. doi:10.1080/00220485.2011.555696.
- *Kerdijk, W., R. A. Tio, B. F. Mulder, and J. Cohen-Schotanus. 2013. "Cumulative Assessment: Strategic Choices to Influence Students' Study Effort." *BMC Medical Education* 13 (172): 1–7. doi:10.1186/1472-6920-13-172.
- *Kibble, J. 2007. "Use of Unsupervised Online Quizzes as Formative Assessment in a Medical Physiology Course: Effects of Incentives on Student Participation and Performance." *Advances in Physiology Education* 31 (3): 253–260. doi:10.1152/advan.00027.2007.
- *Kibble, J. D., T. R. Johnson, M. K. Khalil, L. D. Nelson, G. H. Riggs, J. L. Borrero, and A. F. Payer. 2011. "Insights Gained from the Analysis of Performance and Participation in Online Formative Assessment." *Teaching and Learning in Medicine* 23 (2): 125–129. doi:10.1080/10401334.2011.561687.
- *Kim, M., and J. Ryu. 2013. "The Development and Implementation of a Web-based Formative Peer Assessment System for Enhancing Students' Metacognitive Awareness and Performance in Ill-structured Tasks." *Educational Technology Research and Development* 61 (4): 549–561. doi:10.1007/s11423-012-9266-1.
- *van der Kleij, F. M., T. J. H. M. Eggen, C. F. Timmers, and B. P. Veldkamp. 2012. "Effects of Feedback in a Computer-based Assessment for Learning." *Computers & Education* 58 (1): 263–272. doi:10.1016/j.compedu.2011.07.020.
- *Knight, J. K., and W. B. Wood. 2005. "Teaching More by Lecturing Less." *Cell Biology Education* 4 (4): 298–310. doi:10.1187/05-06-0082.
- *Krasne, S., P. F. Wimmers, A. Relan, and T. A. Drake. 2006. "Differential Effects of Two Types of Formative Assessment in Predicting Performance of First-year Medical Students." *Advances in Health Sciences Education* 11 (2): 155–171. doi:10.1007/s10459-005-5290-9.
- *Krontiris-Litowitz, J. 2008. "Using Truncated Lectures, Conceptual Exercises, and Manipulatives to Improve Learning in the Neuroanatomy Classroom." *Advances in Physiology Education* 32 (2): 152–156. doi:10.1152/advan.00103.2007.
- *Lawrence, N. K. 2013. "Cumulative Exams in the Introductory Psychology Course." *Teaching of Psychology* 40 (1): 15–19. doi:10.1177/00986283112465858.
- Leeming, F. C. 2002. "The Exam-a-day Procedure Improves Performance in Psychology Classes." *Teaching of Psychology* 29 (3): 210–212.
- *Lemus-Zuniga, L. G., J. M. Montanana, F. Buendia-Garcia, J. L. Poza-Lujan, J. L. Posadas-Yaguee, and J. V. Benlloch-Dualde. 2015. "Computer-assisted Method Based on Continuous Feedback to Improve the Academic Achievements of First-year Students on Computer Engineering." *Computer Applications in Engineering Education* 23 (4): 610–620. doi:10.1002/cae.21633.
- *MacDermott, R. J. 2013. "The Impact of Assessment Policy on Learning: Replacement Exams or Grade Dropping." *Journal of Economic Education* 44 (4): 364–371. doi:10.1080/00220485.2013.825114.
- *Maley, M. A. L., J. R. Harvey, W. B. De Boer, N. W. Scott, and G. E. Arena. 2008. "Addressing Current Problems in Teaching Pathology to Medical Students: Blended Learning." *Medical Teacher* 30 (1): E1–E9. doi:10.1080/01421590701753575.
- *Marden, N. Y., L. G. Ulman, F. S. Wilson, and G. M. Velan. 2013. "Online Feedback Assessments in Physiology: Effects on Students' Learning Experiences and Outcomes." *Advances in Physiology Education* 37 (2): 192–200. doi:10.1152/advan.00092.2012.
- *McNulty, J. A., B. R. Spiritu, A. E. Hoyt, D. C. Ensminger, and A. J. Chandrasekhar. 2015. "Associations between Formative Practice Quizzes and Summative Examination Outcomes in a Medical Anatomy Course." *Anatomical Sciences Education* 8 (1): 37–44. doi:10.1002/ase.1442.
- *Micklewright, D., L. Pearsall, M. Sellens, and N. Billam. 2010. "Changes in Approaches to Learning among Undergraduate Sports Science Students following a Programme of Weekly Online Assessments." *Journal of Hospitality Leisure Sport & Tourism Education* 9 (2): 141–155. doi:10.3794/johlste.92.216.
- *Mitra, N. K., and A. Barua. 2015. "Effect of Online Formative Assessment on Summative Performance in Integrated Musculoskeletal System Module." *BMC Medical Education* 15 (1): 29. doi:10.1186/s12909-015-0318-1.
- *Mulder, R., C. Baik, R. Naylor, and J. Pearce. 2014. "How Does Student Peer Review Influence Perceptions, Engagement and Academic Outcomes? A Case Study." *Assessment & Evaluation in Higher Education* 39 (6): 657–677. doi:10.1080/02602938.2013.860421.
- *Munson, A., and R. Pierce. 2015. "Flipping Content to Improve Student Examination Performance in a Pharmacogenomics Course." *American Journal of Pharmaceutical Education* 79 (7): 103. doi: 10.5688/ajpe797103.

- *Nelson, J., D. F. Robison, J. D. Bell, and W. S. Bradshaw. 2009. "Cloning the Professor, an Alternative to Ineffective Teaching in a Large Course." *CBE—Life Sciences Education* 8 (3): 252–263. doi:10.1187/cbe.09-01-0006.
- *Nicol, D. 2009. "Assessment for Learner Self-regulation: Enhancing Achievement in the First Year Using Learning Technologies." *Assessment & Evaluation in Higher Education* 34 (3): 335–352. doi:10.1080/02602930802255139.
- *Olson, B. L., and J. L. McDonald. 2004. "Influence of Online Formative Assessment upon Student Learning in Biomedical Science Courses." *Journal of Dental Education* 68 (6): 656–659.
- *Palmer, E. J., and P. G. Devitt. 2008. "Limitations of Student-driven Formative Assessment in a Clinical Clerkship. A Randomised Controlled Trial." *BMC Medical Education* 8 (1): 29. doi:10.1186/1472-6920-8-29.
- *Palmer, E., and P. Devitt. 2014. "The Assessment of a Structured Online Formative Assessment Program: A Randomised Controlled Trial." *BMC Medical Education* 14 (1): 8. doi:10.1186/1472-6920-14-8.
- *Paschal, C. B. 2002. "Formative Assessment in Physiology Teaching using a Wireless Classroom Communication System." *Advances in Physiology Education* 26 (4): 299–308. doi:10.1152/advan.00030.2002.
- *Perdigones, A., J. L. Garcia, V. Valino, and C. Raposo. 2009. "Assessing Heterogeneous Student Bodies Using a Methodology That Encourages the Acquisition of Skills Valued by Employers." *Assessment & Evaluation in Higher Education* 34 (4): 389–400. doi:10.1080/02602930802071056.
- *Priya, M., M. S. Muthu, D. Amaral, and E. Thomas. 2012. "Continuous Assessment of Undergraduate Students at a Dental College in India." *Journal of Dental Education* 76 (4): 501–508.
- *Ramirez, B. U. 2010. "Effect of Self-assessment on Test Scores: Student Perceptions." *Advances in Physiology Education* 34 (3): 134–136. doi:10.1152/advan.00021.2010.
- *Raupach, T., J. Brown, S. Anders, G. Hasenfuss, and S. Harendza. 2013. "Summative Assessments Are More Powerful Drivers of Student Learning than Resource Intensive Teaching Formats." *BMC Medicine* 11 (1): 61. doi:10.1186/1741-7015-11-61.
- *Rezaei, A. R. 2015. "Frequent Collaborative Quiz Taking and Conceptual Learning." *Active Learning in Higher Education* 16 (3): 187–196. doi:10.1177/1469787415589627.
- Richardson, M., C. Abraham, and R. Bond. 2012. "Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-analysis." *Psychological Bulletin* 138 (2): 353–387. doi:10.1037/a0026838.
- *Sample, S. J., J. A. Bleedorn, S. L. Schaefer, A. Mikla, C. W. Olsen, and P. Muir. 2014. "Students' Perception of Case-based Continuous Assessment and Multiple-choice Assessment in a Small Animal Surgery Course for Veterinary Medical Students." *Veterinary Surgery* 43 (4): 388–399. doi:10.1111/j.1532-950X.2014.12145.x.
- *Sargent, C. S., and A. A. Curcio. 2012. "Empirical Evidence That Formative Assessments Improve Final Exams." *Journal of Legal Education* 61 (3): 379–405.
- *Schlegel, E. F. M., and N. J. Selfridge. 2014. "Fun, Collaboration and Formative Assessment: Skinquization, a Class Wide Gaming Competition in a Medical School with a Large Class." *Medical Teacher* 36 (5): 447–449. doi:10.3109/01421519x.2014.888409.
- Shute, V. J. 2008. "Focus on Formative Feedback." *Review of Educational Research* 78 (1): 153–189. doi:10.3102/0034654307313795.
- *Siweya, H. J., and P. Letsoalo. 2014. "Formative Assessment by First-year Chemistry Students as Predictor of Success in Summative Assessment at a South African University." *Chemistry Education Research and Practice* 15 (4): 541–549. doi:10.1039/c4rp00032c.
- *Snowball, J. D., and M. Mostert. 2013. "Dancing with the Devil: Formative Peer Assessment and Academic Performance." *Higher Education Research & Development* 32 (4): 646–659. doi:10.1080/07294360.2012.705262.
- *Stellmack, M. A., N. K. Keenan, R. R. Sandidge, A. L. Sippl, and Y. L. Konheim-Kalkstein. 2012. "Review, Revise, and Resubmit: The Effects of Self-critique, Peer Review, and Instructor Feedback on Student Writing." *Teaching of Psychology* 39 (4): 235–244. doi:10.1177/0098628312456589.
- *Sullivan, P. B., N. Gregg, E. Adams, C. Rodgers, and J. Hull. 2013. "How Much of the Paediatric Core Curriculum Do Medical Students Remember?" *Advances in Health Sciences Education* 18 (3): 365–373. doi:10.1007/s10459-012-9375-y.
- *Taradi, S. K., M. Taradi, and K. I. Radic. 2005. "Integration of Online Formative Assessments into Medical Education: Experience from University of Zagreb Medical School, Croatia." *National Medical Journal of India* 18 (1): 39–40.
- *Thome, G., H. Hovenberg, and G. Edgren. 2006. "Portfolio as a Method for Continuous Assessment in an Undergraduate Health Education Programme." *Medical Teacher* 28 (6): E171–E176. doi:10.1080/01421590600776511.
- Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276. doi:10.3102/00346543068003249.
- *Tuunila, R., and M. Pulkkinen. 2015. "Effect of Continuous Assessment on Learning Outcomes on Two Chemical Engineering Courses: Case Study." *European Journal of Engineering Education* 40 (6): 671–682. doi:10.1080/03043797.2014.1001819.
- *Velan, G. M., P. Jones, H. P. McNeil, and R. K. Kumar. 2008. "Integrated Online Formative Assessments in the Biomedical Sciences for Medical Students: Benefits for Learning." *BMC Medical Education* 8 (1): 52. doi:10.1186/1472-6920-8-52.
- *Wakefield, C., J. Adie, E. Pitt, and T. Owens. 2014. "Feeding Forward from Summative Assessment: The Essay Feedback Checklist as a Learning Tool." *Assessment & Evaluation in Higher Education* 39 (2): 253–262. doi:10.1080/02602938.2013.822845.
- *Williams, N. A., W. Bland, and G. Christie. 2008. "Improving Student Achievement and Satisfaction by Adopting a Blended Learning Approach to Inorganic Chemistry." *Chemistry Education Research and Practice* 9 (1): 43–50. doi:10.1039/b801290n.

- *Wilson, K., C. Boyd, L. Chen, and S. Jamal. 2011. "Improving Student Performance in a First-year Geography Course: Examining the Importance of Computer-assisted Formative Assessment." *Computers & Education* 57 (2): 1493–1500. doi:[10.1016/j.compedu.2011.02.011](https://doi.org/10.1016/j.compedu.2011.02.011).
- *Wininger, S. R. 2005. "Using Your Tests to Teach: Formative Summative Assessment." *Teaching of Psychology* 32 (3): 164–166. doi:[10.1207/s15328023top3203_7](https://doi.org/10.1207/s15328023top3203_7).
- *Zhang, N., and C. N. R. Henderson. 2015. "Can Formative Quizzes Predict or Improve Summative Exam Performance?" *Journal of Chiropractic Education* 29 (1): 16–21. doi:[10.7899/jce-14-12](https://doi.org/10.7899/jce-14-12).
- van Zundert, M., D. Sluijsmans, and J. van Merriënboer. 2010. "Effective Peer Assessment Processes: Research Findings and Future Directions." *Learning and Instruction* 20 (4): 270–279. doi:[10.1016/j.learninstruc.2009.08.004](https://doi.org/10.1016/j.learninstruc.2009.08.004).