

# Chapter 4

## Inspiring Transformation Through TESTA's Programme Approach

Tansy Jessop

**Abstract** This chapter explores evidence from 75 undergraduate degree programmes at 14 UK universities about students' experience of assessment and feedback. The data was collected through the *Transforming the Experience of Students Through Assessment* (TESTA) project. TESTA illuminates students' whole programme experience of assessment and feedback within the context of modular curriculum design. The methodology consists of an audit to ascertain dimensions of the assessment environment and focus groups with students. Analysis explores the relationship between assessment design and students' lived experience. Findings show the prevalence of high summative and low formative assessment diets and disconnected feedback which students find difficult to use. High summative diets reinforce students' instrumental approach to learning. A lack of formative assessment impacts on engagement in learning, diminishing opportunities for risk-taking, creativity and wider reading. Findings from TESTA have prompted educationally principled strategies: rebalancing the number of summative and formative tasks with greater connections between them, devising formative assessment tasks valued by both students and staff and designing feedback to feedforward. Wider implications include establishing institutional mechanisms to ensure principled, evidence-based and programme-focused assessment and feedback design within existing curriculum processes.

### Introduction

This chapter is based on triangulated data about programme assessment from the *Transforming the Experience of Students Through Assessment* (TESTA) project (TESTA, 2015). TESTA started as a 3-year funded UK Higher Education Academy (HEA) National Teaching Fellowship Project (2009–2012) to investigate programme-level assessment on seven programmes in four similar small universities. The purpose of the research was to explore the impact of modular

---

T. Jessop (✉)  
Southampton Solent University, Southampton, UK  
e-mail: [tansy.jessop@solent.ac.uk](mailto:tansy.jessop@solent.ac.uk)

systems on assessment and feedback design and consequently on how students learn. By 2016, 4 years after funding ceased, more than 50 universities in the UK and universities in Australia, India, Canada, the USA and South Africa have drawn on TESTA's approach. This is testimony to the sector's appetite for understanding whole programme assessment patterns, but also to the value of TESTA's research methodology and its participatory change process.

Modular assessment has come under criticism for fragmenting the curriculum, fostering a lack of connection and coherence and blurring chronological progression. The introduction of semesters and modular curricula has compressed learning into short, contained units to the extent that 'slow learning' and formative assessment are squeezed out (Harland, McLean, Wass, Miller, & Sim, 2014; Knight, 2001; Knight & Yorke, 2003; Rust, 2000). TESTA provides an evidence base which gives insight into the impact of modular assessment environments on student learning. It enables programme teams to redesign assessment and feedback holistically, with articulation between the evidence, assessment principles and quality assurance frameworks (Gibbs & Dunbar-Goddet, 2007, 2009; Jessop, El Hakim, & Gibbs, 2011, 2014; Jessop, McNab, & Gubby, 2012). TESTA's approach has brought credible evidence to sector-wide discussions about developing a strategic focus on programme assessment and feedback (Bloxham & Boyd, 2007; Knight, 2000; PASS, 2009–2012).

The following sections outline TESTA's research methodology and explore key findings, with examples of best practice arising from undertaking the change process. It concludes with a strategy for scaling up assessment transformation institutionally.

## Research Methods

Previous studies have used various combinations of the three TESTA methods to provide a critical perspective of degree programme environments. A previously published TESTA study triangulated the audit, the Assessment Experience Questionnaire (AEQ) and focus group data to provide an analysis of programme assessment environments (Jessop, El Hakim, & Gibbs, 2013). A follow-up study triangulated audit and AEQ data to illuminate disciplinary assessment practices (Jessop & Maleckar, 2014). Since the publication of these studies, based on smaller data sets (23 and 18 programmes, respectively), many more programmes have undertaken TESTA across the UK. This chapter analyses data through the lens of two of the three TESTA methods, namely, the audit and focus groups. The rationale is to add new knowledge based on a larger sample than previous studies which relied heavily on statistical methods (Jessop et al., 2013, 2014). This chapter gives a distinctive qualitative perspective on programme assessment environments using large-scale data, by triangulating the hard count data of the TESTA audit with student voice data from focus groups.

The TESTA audit distinguishes elements of the assessment environment (Gibbs & Dunbar-Goddet, 2009). It consists of discussion with the programme leader over course documents to map the key features of assessment and feedback over 3 years of an undergraduate programme (Jessop, 2010a). These features are:

- Number of summative assessments
- Number of formative assessment tasks
- Number of different varieties of assessment types
- Proportion of the assessment diet by examinations
- Time it takes to return marks and feedback
- Amount of written feedback over 3 years
- Amount of oral feedback over 3 years

The programme audit represents the 'planned curriculum' (Stenhouse, 1975), providing hard count data about key features of students' experience over 3 years of an undergraduate degree. In most UK universities, the definitive documents of a programme normally undergo revision and scrutiny every 5 or 6 years, with modules, assessments, learning outcomes and the content of a programme being revised in the light of institutional, staffing and subject developments. This process, known as periodic review, provides programmes with external and internal scrutiny to assure universities and students of the quality of provision. While this documentation is publically available, the TESTA audit brings to light the less visible aspects of a programme which may not be contained in the documents, for example, formative assessment tasks. The audit is a mixed methods approach drawing on documentary and interview evidence.

Focus groups provide the second vein of data for this study. The hallmark of focus groups is their 'explicit use of group interaction to produce data and insights that would be less accessible without the interaction found in the group' (Morgan, 1997, p. 2). TESTA focus groups prompt discussion on four themes: the assessment diet, feedback, the influence of assessment on study habits and students' understanding of goals and standards (Jessop, 2010b). These themes are hospitable and allow students to discuss wider issues about assessment and feedback on their programme. Typically a focus group consists of an hour-long discussion between a group of between five and eight final year students, facilitated by a researcher.

Sampling of programmes varies institutionally. At Winchester, TESTA began as a voluntary exercise which programmes signed up for out of interest or the desire to enhance their programme's assessment. Since 2013, TESTA has been embedded in cyclical quality processes, so that all undergraduate programmes undergoing six yearly periodic reviews are required to participate in TESTA. This evidence gathering and planning informs curriculum design. Other institutions have nominated programmes to participate; alternatively whole departments, faculties and programmes have signed up (or been signed up) to participate in TESTA.

Data analysis of the audit is a relatively simple process of transposing the data, recorded on a flipchart, into a brief report which summarises occurrences of assessment, and volumes of feedback, for example. The rigour of the audit is ensured through investigator triangulation (Cohen, Manion, & Morrison, 2007) when more

than one researcher is eliciting the data. More often, the accuracy of interpretation is ensured through ‘member-checking’ (Lincoln & Guba, 1985) which involves the draft audit document being sent to the programme leader to check the accuracy of information and interpretation.

Focus group data are recorded and transcribed. The transcripts are uploaded into qualitative data analysis software (Atlas.ti), for the purpose of thematic analysis. The researcher codes units of meaning, using either thematic coding (drawing on existing constructs about how students learn from feedback) or generative coding (open to new constructs from the data). An example of a thematic code might be ‘confusing criteria’, with criteria being an explicit area of investigation. In contrast, a generative code might be ‘marker variation’, which students raise as an implicit barrier to understanding goals and standards. The levels of codes also vary between concrete descriptions to more abstract notions such as strategic or instrumental approaches to learning. There is a guide to coding focus group data on the TESTA website (Jessop, 2011).

Normally, the TESTA process consists of representing audit, Assessment Experience Questionnaire and focus group data in a case study which is discussed with programme teams. The case study is the focal point for developing strategies to enhance curriculum design and pedagogic practice. In this chapter, the key findings draw only on audit and focus group data. The audit data provide a summary of the numbers/proportions of certain assessment and feedback activities, in ranges and medians, across the whole sample. As in any qualitative research with large volumes of textual data, coding segments of data underpins the process of developing themes. The key findings in this chapter represent recurrent themes, derived from a systematic qualitative analysis of transcripts from participating programmes.

## Key Findings

Three main themes in TESTA data are now explored. The first is the variation in assessment patterns and implications for student learning. The second demonstrates the prevalence of high summative assessment diets in contrast to low occurrences of formative assessment. Thirdly, the data sheds light on episodic and haphazard feedback which does not connect to the next task or across modules. These themes all impact on the student learning experience, contributing either to surface or deep learning (Marton & Saljo, 1976), or strategic behaviour (Miller & Parlett, 1974). Student alienation, characterised by ‘playing the game’ in a performative way, arises partly from flaws in the design of assessment and feedback (Boud & Molloy, 2013; Mann, 2001; Miller & Parlett, 1974).

Before discussing the themes, it is worth clarifying how the terms formative and summative assessment are used in this chapter. TESTA defines summative assessment as tasks which are graded and count towards the degree either as pass/fail or as grades; in contrast, formative tasks do not count towards the degree, are required to be done by all students and elicit feedback. Formative assessment has been described as a ‘fuzzy concept’, and its elusiveness has stimulated much debate

and contestation (Taras, 2008; Torrance, 2012; Yorke, 2003). While recognising the problematic nature of defining formative and summative assessment, TESTA adheres to Shepard's distinction between summative and formative: 'summative assessment measures students' achievement by a grade, while formative gives qualitative insights about students' understandings and misconceptions to improve learning' (Shepard, 2005). This distinction is founded on the belief that formative generally plays a different role to summative, privileging reflection and action on feedback, and orienting students towards future performance. In contrast, summative assessment often occurs at the end of modules and orients students towards grades rather than future performance, which may occur in an unrelated area of study, albeit the same discipline.

The key findings section expands on each theme using data from TESTA. Following each theme, there are examples of best practice drawn from programmes and institutions which have engaged with the change process in TESTA. Strategies to implement best practice emerge from a rich discussion of the data in the light of assessment principles.

## Theme 1: Variations in Assessment Environments

Undergraduate degree programmes demonstrate extreme variations in their assessment environments. The disciplines represented among the 75 within the sample include:

- Pure sciences (e.g. mathematics, chemistry)
- Applied sciences (e.g. engineering, pharmacy)
- Humanities and social sciences (e.g. history, sociology)
- Applied 'soft' disciplines (e.g. education, social work)
- Creative subjects (e.g. drama, dance, creative writing)

The following table shows the range and medians of TESTA audit data across a 3-year programme of study within the sample of 75 programmes (Table 4.1).

**Table 4.1** Ranges and medians for TESTA audit data (n=75 programmes)

Characteristic	Range	Median
Number of summative assessments	12–227	43
Number of formative assessments	0–116	5
Varieties of assessment	5–21	11
Proportion of tasks by exam	0–87 %	20 %
Time to return marks and feedback	10–42 days	21 days
Volume of oral feedback	37–1800 min	282 min
Volume of written feedback	936–22,000	6165 words

Variations may occur because of different ‘ways of thinking and practicing’ in the disciplines (Hounsell & Anderson, 2009). Research on ‘signature pedagogies’ (Shulman, 2005) illustrates the ways in which different professions and disciplines enact their knowledge, shown through ‘idiosyncratic organisation, set of artefacts, assumptions and practices peculiar to learning and teaching in the discipline’ (Donald, 2009, p.40). In the sciences, for example, assessment patterns often contain small and frequent formative tasks, partly to ensure that students master concepts incrementally, and to avoid gaps in understanding linked to the next concept (Jessop & Maleckar, 2014).

The curriculum design process also contributes to variations. Aside from more tightly regulated professional programmes, most lecturers exert a significant degree of autonomy in the design of assessment tasks on modules (Bridges, 2000). Programmes are commonly assembled module-by-module according to the content being covered, without an eye on the whole programme’s assessment design. Module leaders often design assessment tasks in isolation and without connection to the wider programme (Jessop et al., 2012).

TESTA has demonstrated that variations in assessment environments exist on a scale that requires attention to ensure comparability in the student learning experience. The most striking variations occur in the number of summative and formative assessments, the proportion of exams and the amount of feedback students typically receive. Variations are likely to influence study behaviours. For example, a high proportion of summative assessment with low formative has been shown to lead to narrowly focused effort, strategic behaviour and a lack of deep learning, evidenced in focus group comments and by low scores on the Quantity of Effort and Deep Learning Scales on the AEQ (Jessop et al., 2013). Many small and frequent summative assessment tasks foster instrumental and grade-conscious approaches to learning (Harland et al., 2014). Bite-sized small tasks may also lead to superficial learning as students are not challenged to undertake large, independent tasks which integrate and connect learning from across the programme (Ashford-Rowe, Herrington, & Brown, 2013; Harland, McLean, Wass, Miller, & Sim, 2014, Jessop et al., 2013). Students also describe incidences of marker variation, which are evidenced in low scores on the Clear Goals and Standards Scale on the AEQ (see Jessop et al., 2013).

TESTA has helped institutions and programmes to address variations in practice within individual universities. The following case studies illustrate actions taken to reduce variations, in the interests of parity of the student experience:

### **Case Study 1: Addressing variations at the institutional level**

*The Problem:* Students experience widely differing assessment and feedback practice, particularly with some programmes having extremely high summative demands alongside invisible, uneven or non-existent formative tasks.

(continued)

There is a huge variety of randomly sequenced types of assessment across the degree with confusing demands for students.

*Strategy:* All programmes go through TESTA at periodic review. TESTA demonstrates how many summative assessment tasks a student will experience, a figure which is not self-evident from the modular nature of the documents. The concept of a reasonable programme assessment load is discussed against the backdrop of sector-wide data and assessment principles, with the aim of rebalancing the ratio of summative and formative. Appropriate and well-designed formative assessments are written into the documentation and planned to articulate with summative tasks. Varieties of assessment are sequenced through the degree, in many cases streamlining these so as to enable a coherent journey of learning through the degree.

### **Case Study 2: Variations within a programme**

*The Problem:* Students describe wide variations between markers on a programme, characterised by some lecturers marking harshly, others more generously. Student focus groups evidence perceptions of bias, varying styles of marking, different approaches and standards.

*Strategy: Calibration of standards*

Facilitate a programme calibration exercise involving all team members. Begin by discussing criteria using flipchart and discussion; compare with existing written criteria and discuss tacit criteria and standards. Mark two or three anonymous written pieces. Collect marker's individually assigned marks before a round-table discussion. This negates programme team power relations and hierarchies, ensuring that markers commit to their marks before discussing why they have assigned them. Display the range of marks. Discuss marks in relation to the agreed criteria and come to consensus. The discussion should begin to articulate more of a common standard. Repeat with a different piece. Repeat calibration annually.

*Note:* Calibration is different from moderation. It is a much more all round, relaxed and open discussion of team standards. It is not linked to time-bound marking processes and formal institutional quality assurance processes.

## **Theme 2: High Summative and Low Formative Assessment**

TESTA audit data shows that students typically encounter 43 summative tasks over 3 years, while experiencing only five formative tasks. This is a ratio of 8:1 of summative to formative tasks. In contexts of high summative assessment, students frequently encounter many, bite-sized summative tasks, which may distribute effort

but not challenge them; these students often become strategic performers in an ‘assessment arms race’ (Harland et al., 2014), depleting their capacity for deep learning. On the surface, they will have worked harder than other students, but their learning is likely to have been trivialised by small, frequent and narrowly focused tasks. In focus groups, students describe how summative assessment dominates their study behaviour and narrows their focus. Below are quotations which exemplify these issues:

A: A lot of people don’t do wider reading. You just focus on your essay question.

B: I always find myself going to the library and going ‘These are the books related to this essay’ and that’s it (Archaeology).

If someone said what did you learn on your degree, I’d basically sum it up as saying I learnt what I needed to learn for assignments; I learnt what I learnt because I needed to complete an assignment, rather than I learnt because I was really interested in the whole thing (English Language Studies).

In Weeks 9 to 12 there is hardly anyone in our lectures because we’re too stressed. I’d rather use those two hours of lectures to get the assignment done (Theology and Religion).

While summative assessment tasks may be learning oriented (Carless, 2007, 2015), TESTA data shows that this is difficult to achieve across compartmentalised modules designed in an atomised way. Students describe the timing and volume of assessment as interfering with learning-oriented assessment, especially when too much assessment crowds out reflection and fosters a grade-driven approach (Knight & Yorke, 2003; Lizzio, Wilson, & Simons, 2002). In the context of high summative diets, an instrumental culture flourishes, because students focus on achieving grades in the shadow of ever-present deadlines (Harland et al., 2014; Jessop et al., 2013). Modular design has multiplied the number of assessment tasks and disconnected them from one another across modules. Students indicate how stressful and demotivating a succession of summative tasks can be, replacing assessment *of* learning with assessment *as* learning and bypassing assessment *for* learning (Torrance, 2007):

The quantity of assessed work is very tiring. We’d rather genuinely study the subject (Education).

It’s been non-stop assignments, and I’m now free of assignments until the exams – I’ve had to rush every piece of work I’ve done (History).

There was a full two weeks of madness, because there was the poster submission, thesis, then the poster presentation, then the exam. It was a very stressful period at the time. Motivation was hard to come by (Pharmacy).

Focus group and audit data imply an unspoken agreement between academics and students that summative assessment is the main way to drive student effort. Without the incentive of a grade, some students say that they are disinclined to undertake academic work. As one student wryly observed, ‘The lecturers have the problem of actually getting the students to go away and learn. If it is not being assessed, if there’re no actual consequences of not doing it, most students are going to sit in the bar’ (Computing Student). The audit data bears out a tentative embrace of formative tasks in programme design, with lecturers underplaying it to the extent that one



in five programmes contains zero formative assessment, and students typically experience less than two formative tasks each year.

Extremely low formative assessment occurs in spite of overwhelming evidence in the literature of its effectiveness in helping students to learn from assessment. Black and Wiliam's large-scale analysis of factors influencing learning concluded that 'innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains' (Black & Wiliam, 1998, p. 40). The reasons for formative assessment's effectiveness include its capacity for 'short-circuiting the randomness and inefficiency of trial-and-error learning' (Sadler, 1989, p. 120), and its capacity to help students fine-tune their work by coming to a deeper understanding of goals and standards (Boud, 2000; Nicol & Macfarlane-Dick, 2006).

Students prioritise assessment tasks which count towards their degrees. In the context of high summative assessment demands, it is unsurprising that formative tasks are undervalued when they compete for time and effort with tasks which count. The combination of high summative demands on concurrent modules, and short semesters, make it almost inevitable that formative tasks are squeezed out. Comments from students in focus groups evidence these competing priorities:

It didn't count for anything, so if you didn't do it, it didn't matter (Mathematics).

It's a little bit pointless for me because I'd rather put all my energy and efforts into marked ones and get really high grades on them and not bother with the others (Philosophy).

What is the point of putting that much effort in when I have this much time to do an assessment that counts towards my degree? I find it really frustrating when people ask for ten page reports and presentations which don't count and I am thinking why am I doing this?! It's brilliant practice but... (Business and Management).

The low value of formative is compounded by issues with the distribution of assessment in compact semesters. Most programmes have two assessment points per module, which cluster at the mid- and end points of modules. The timing of formative tasks is a complex and important aspect of design, influencing students' capacity to engage with formative tasks.

The following case studies demonstrate actions which programme teams have taken to address the challenges of high summative and low formative assessment diets:

### **Case Study 1: Rebalancing summative and formative**

*Problem:* TESTA uncovers that several programmes in a business school have a typical assessment load of 48 summative tasks. Students are working for each assessment *as learning*, ignoring wider reading and set tasks.

*Strategy:* Departmental decision is made by the head of school to revalidate all programmes with mandated limits of one summative assessment on each module and three formative tasks leading up to the summative. Timing and sequencing of all assessments are agreed to prevent competing clashes. All

(continued)

programmes move to a summative assessment load of 24 tasks and a ratio of 3:1 of formative to summative.

*Design complexities:* There is variation in the success of the formative design because some formative ‘teaches to the test’, with the result that students fine-tune similar work to the summative to achieve better grades. Strategies are put in place to support lecturers in the design of formative which synchronises with summative but is challenging and stand-alone, yet conceptually linked to the summative.

*Feedback complexities:* Formative work is often peer reviewed in class, so that lecturers do not increase marking loads by having four marking occurrences ( $3 \times$  formative plus  $1 \times$  summative) instead of two as in the past ( $2 \times$  summative). Best practice ensures discussion and dialogue about marking and student use or co-creation of criteria.

### **Case Study 2: Students as producers of formative work**

*Problem:* TESTA shows ineffective formative tasks which are only done by keen students or ‘dashed off’ by students who do not value formative work, particularly when it competes with summative demands on other modules.

*Strategy:* Whole programme strategy is adopted to reduce summative assessment load on all modules. Well-designed formative tasks command student attention and interest, replacing certain summative tasks. Principles of good formative assessment practised as a result of undertaking TESTA include as follows:

- (a) Public facing so as motivate students to perform and contribute to knowledge generation.
- (b) Links to the summative task in a challenging and enriching way.
- (c) Builds in elements of collaboration and accountability.
- (d) Involves challenging research, theory or project design.
- (e) Links to the discipline and has a genuine purpose.
- (f) Encourages students to be creative and take risks.
- (g) Engenders feedback from peers or the tutor and encourages reflection and inner dialogue.
- (h) Students are required to undertake it as a gateway to completing the summative task.

#### *Idea 1: Blogging as formative*

Students blog fortnightly in class on academic readings. In alternate weeks, students read several fellow students’ blogs and spend the in-class hour commenting on posts. The summative assessment is constructed around

(continued)

conceptual understandings developed through the blogging or may be a synthesis of different arguments and positions. 'Think aloud' data on blogging shows that it prompts engagement with academic texts, reflection, distribution of effort and deep learning (Jessop, 2015).

*Idea 2: Project development and design*

Assessment consists of large, challenging, collaborative (or individual) projects which involve designing publically available artefacts, such as films or posters or a publically disseminated research project, to address an issue or problem in the discipline. In one case, the team of lecturers split a 12-week module into 6 × weeks of lectures and 6 × weeks of collaborative project work, guided in class. Students produced media artefacts which were showcased in the final weeks.

### Theme 3: Disconnected Feedback

Written feedback volumes are calculated by sampling in-text and summary feedback from each cohort, from a variety of different markers. Typically, students receive around 7400 words of written feedback over a 3-year undergraduate degree. This represents a significant amount of time, thought and effort in crafting feedback, yet in the National Student Survey (NSS), students rate the effectiveness of feedback lower than any other aspect of teaching. Question 9 on the NSS 'Feedback on my work has helped me clarify things I did not understand' is consistently the lowest scored of the assessment and feedback questions. In England, only 66 % (2014) and 67 % (2015) of full-time undergraduates on taught courses affirmed that feedback helped them to clarify things they did not understand (HEFCE, 2015). At least one third of feedback is not working.

Evidence from TESTA gives some explanation for the broken state of feedback. It sheds light on the waste of resources when feedback ends in a cul-de-sac. Feedback which is end loaded, occurring after the module has finished, has little chance of improving student performance. Audit and focus group data show a lack of articulation of feedback across modules, with few opportunities for feeding forward designed into the process. One-off, episodic and piecemeal feedback which 'dangles the data' (Sadler, 1989) has questionable value when it is not designed as a process which helps students to reflect and act on it with a future orientation (Boud & Molloy, 2012, 2013). In a modular system, students are more inclined to compartmentalise their learning and ignore feedback which does not have much chance of feeding forward, as these comments demonstrate:

It's difficult because your assignments are so detached from the next one you do for that subject. They don't relate to each other (Media Studies).

It is so dependent on whether it is the first assignment for the module or not... say you have a fifteen credit module where you have two essays and it's the first essay back,

then you would probably take quite a lot on board but then for the essay after that I was quite pleased with my marks so I just was like fine! As bad as it sounds, the module is over (Business).

Because it's at the end of the module, it doesn't feed into our future work (History).

The feedback is generally focused on the module (Primary Education).

I read it and think 'Well, that's fine but I've already handed it in now and got the mark. It's too late' (Creative Writing).

Through doing TESTA, programme teams have developed strategies for making feedback connect across modules, exemplified below:

### **Case Study 1: Building reflection into feedback in a way which leads to action**

*Problem:* Students do not make use of end-of-module feedback unless there is a problem with their mark.

*Strategy:* Programmes give students feedback on ways to improve their work, which students are required to reflect on in writing in their next task. The lecturer will not mark the next task unless students have shown how they have addressed previous feedback.

### **Case Study 2: Building dialogue into feedback in a way which leads to action**

*Problem:* Students feel that they are the passive victims of feedback which is a one-way transaction from expert to novice.

*Strategy 1:* Students open the conversation by indicating what they would like feedback on, for example, targeted towards what they feel they understood or argued well or where they feel conceptually fuzzy. Markers respond to the conversation with feedback. This is dialogue!

*Strategy 2:* Markers release comments only. Students need to write a brief response to feedback and attend a tutorial to receive their mark. Students have reflected on and engaged in dialogue about their feedback without being distracted by the mark.

*Strategy 3:* Markers give generic feedback about strengths, weaknesses and issues based on a sample of marking very soon after the hand-in date. This helps students to think through their own practice from memory, prompting inner dialogue.

### **Case Study 3: Integrated assessment across modules**

*Problem:* Students make little use of end-of-module feedback because it is too late, or modules are viewed as 'finished business'.

(continued)

*Strategy:* Curriculum design makes more explicit links and connections across modules through a capstone assessment which threads across several modules. This may be a reflective portfolio, a group project or a research paper.

## Scaling Up Transformation: An Embedded Institutional Approach

TESTA provides an example of a systematic approach to transforming the student experience of assessment and feedback. It is based on reconceptualising assessment and feedback design as integrated, connected and sequential within curriculum design processes *at the programme level*. There are four key dimensions to bringing about institutional transformation in assessment and feedback to improve student learning, developed through 6 years of implementing TESTA's research and change process. These are:

- Taking a whole programme view of assessment and feedback
- Using an evidence-led approach with a strong element of student voice
- Putting the evidence and principles in the hands of teams to make changes
- Adopting a systemic approach through quality assurance processes

A whole programme approach to assessment and feedback prompts two shifts in perception: firstly, lecturers shift from thinking about curriculum and pedagogy from the silo perspective of 'my' module to a connected view of 'our' programme. Secondly, lecturers start to understand what learning looks like from the programme-wide experience of a student juggling modules and deadlines, in contrast to the funnelled view of seeing students as participants on 'my' module. The programme view enables a shift from a teacher-centred to a student-centred paradigm of learning from assessment and feedback.

The second dimension is TESTA's robust evidence-base, triangulating audit, survey and focus group data in rich and textured case studies. Programme teams describe the data as plausible and compelling. It validates hunches and intuitions about the assessment environment by bringing particular evidence into focus. The evidence may be challenging and discomfiting, but teams are able to engage with it because of its externality, its robustness and its genuine enhancement focus. The way the case studies are constructed ensures that the student voice data is prominent, rather than wrapped up in authorial interpretation (Jessop, 2013; Richardson, 1990).

The third important element in scaling up institutional change through TESTA has been placing the evidence in the hands of the team in order for them to develop holistic curriculum and pedagogic strategies. This respects the autonomy, agency and disciplinary knowledge of programme teams. Meeting over evidence takes the

shape of a consultancy briefing, with rich discussion, contestation, new ideas and ‘what if’ questions being raised. It is fertile ground on which to base curriculum design and pedagogic decisions.

The final dimension of assessment transformation using TESTA has been to scale it up institutionally, by embedding it in quality assurance processes. Systematic embedding of TESTA in cyclical periodic review contributes to evidence-led curriculum design on all revalidating degrees. This model drives institution-wide and programme-specific enhancement effects at the Universities of Winchester and Dundee, with further universities exploring its use. As more programmes engage at Winchester, deans of faculty describe a ‘TESTA effect’: a new appreciation of formative assessment, coupled with principled and evidence-based assessment design. The chair of the UK’s assessment in HE conference has described TESTA as ‘the only thing I can find that seems to be making systematic headway beyond the individual module’ (S. Bloxham, personal communication, 17 March, 2016). Isolating reasons for TESTA’s systematic headway is complex, but two aspects stand out: the focus on the programme and the marriage between an enhancement project and a quality assurance process. These signify TESTA’s potential to impact the institutional assessment culture, in particular, its capacity to create a powerful learning experience for students through systematic transformation.

## References

- Ashford-Rowe, K., Herrington, J., & Brown, C. (2013). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205–222.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Bloxham, S., & Boyd, P. (2007). Planning a programme assessment strategy. In *Developing effective assessment in higher education* (pp. 157–175). Maidenhead, UK: Open University Press.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167.
- Boud, D., & Molloy, E. (Eds.). (2012). *Feedback in higher and professional education: Understanding it and doing it well*. New York: Routledge.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712.
- Bridges, D. (2000). Back to the future: The higher education curriculum in the 21st century. *Cambridge Journal of Education*, 30(1), 37–55.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66.
- Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. London: Routledge.
- Cohen, P. L., Manion, L., & Morrison, K. (2007). *Research methods in education*. New York: Routledge.
- Donald, J. G. (2009). The Commons: Disciplinary and interdisciplinary encounters. In C. Kreber (Ed.), *The university and its disciplines: Teaching and learning within and beyond disciplinary boundaries* (pp. 35–49). New York: Routledge.

- Gibbs, G., & Dunbar-Goddet, H. (2007). *The effects of programme assessment environments on student learning*. Higher Education Academy Report. Retrieved from [https://www.heacademy.ac.uk/sites/default/files/gibbs\\_0506.pdf](https://www.heacademy.ac.uk/sites/default/files/gibbs_0506.pdf)
- Gibbs, G., & Dunbar-Goddet, H. (2009). Characterising programme-level assessment environments that support learning. *Assessment & Evaluation in Higher Education*, 34(4), 481–489.
- Harland, T., McLean, A., Wass, R., Miller, E., & Sim, K. N. (2014). 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.
- Higher Education Funding Council for England (HEFCE). (2015). *2015 – Higher education funding council for England*. Retrieved from <http://www.hefce.ac.uk/lt/nss/results/2015/>
- Hounsell, D., & Anderson, C. (2009). Ways of thinking and practising in biology and history: Disciplinary aspects of teaching and learning environments. In C. Kreber (Ed.), *The university and its disciplines: Teaching and learning within and beyond disciplinary boundaries*. New York: Routledge.
- Jessop, T. (2010a). *Ten steps to auditing a programme*. Retrieved from <http://www.testa.ac.uk/index.php/resources/research-tool-kits/category/11-researchtoolkits>
- Jessop, T. (2010b). *Focus group questions*. Retrieved from <http://www.testa.ac.uk/index.php/resources/research-tool-kits/category/11-researchtoolkits>
- Jessop, T. (2011). *Coding focus group data*. Retrieved from <http://www.testa.ac.uk/index.php/resources/research-tool-kits/category/11-researchtoolkits>
- Jessop, T. (2013). *TESTA case studies*. Retrieved from <http://www.testa.ac.uk/index.php/resources/case-studies/category/8-case-studies>
- Jessop, T. (2015, December). *Flipping the technology-pedagogy equation: Principles to improve assessment and feedback*. Seminar presentation at the University of Sussex, Brighton, UK.
- Jessop, T., El Hakim, Y., & Gibbs, G. (2011). TESTA: Research inspiring change. *Educational Developments*, 12(4), 12–16.
- Jessop, T., El Hakim, Y., & Gibbs, G. (2013). The whole is greater than the sum of its parts: A large-scale study of students' learning in response to different programme assessment patterns. *Assessment & Evaluation in Higher Education*, 39(1), 73–88.
- Jessop, T., El Hakim, Y., & Gibbs, G. (2014). TESTA 2014: A way of thinking about assessment and feedback. *Educational Developments*, 15(2), 21–23.
- Jessop, T., & Maleckar, B. (2014). The influence of disciplinary assessment patterns on student learning: A comparative study. *Studies in Higher Education*, 41(4), 696–711.
- Jessop, T., McNab, N., & Gubby, L. (2012). Mind the gap: An analysis of how quality assurance procedures influence programme assessment patterns. *Active Learning in Higher Education*, 13(3), 143–154.
- Knight, P. T. (2000). The value of a programme-wide approach to assessment. *Assessment & Evaluation in Higher Education*, 25(3), 237–251.
- Knight, P. T. (2001). Complexity and curriculum: A process approach to curriculum-making. *Teaching in Higher Education*, 6(3), 369–381.
- Knight, P. T., & Yorke, M. (2003). *Assessment learning and employability*. Maidenhead, UK: Society for Research into Higher Education & Open University Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
- Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education*, 27(1), 27–52.
- Mann, S. J. (2001). Alternative perspectives on the student experience: Alienation and engagement. *Studies in Higher Education*, 26(1), 7–19.
- Marton, F., & Saljo, R. (1976). On qualitative differences in learning: Outcome and process. *British Journal of Educational Psychology*, 46(1), 4–11.
- Miller, C. M. L., & Parlett, M. R. (1974). *Up to the mark: A study of the examination game*. London: Society for Research into Higher Education.
- Morgan, D. L. (1997). *Focus groups as qualitative research*. London: Sage Publications.

- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.
- Programme Assessment Strategies Project (PASS). (2009–2012). *University of Bradford*. Retrieved from <http://pass.brad.ac.uk>
- Richardson, L. W. (1990). *Writing strategies: Reaching diverse audiences*. Newbury Park, CA: Sage Publications.
- Rust, C. (2000). A possible student-centred assessment solution to some of the current problems of modular degree programmes. *Active Learning in Higher Education, 1*, 126–131.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.
- Shepard, L. A. (2005, October). *Formative assessment: Caveat emptor*. Paper presented at ‘The Future of Assessment: Shaping Teaching and Learning’ Conference, New York, NY.
- Shulman, L. S. (2005). Signature pedagogies in the professions. *Daedalus, 134*(3), 52–59.
- Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann Educational Books.
- Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active Learning in Higher Education, 9*(2), 172–192.
- Transforming the Experience of Students through Assessment (TESTA). (2015). The University of Winchester. Retrieved from <http://www.testa.ac.uk>
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice, 14*(3), 281–294.
- Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education, 38*(3), 323–342.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477–501.