University of Cumbria

Conceptual Acrobatics: Talking about Assessment Standards in the Transparency Era

Journal: Studies in Higher Education

Corresponding Author:

Jane Hudson, ASKe Pedagogy Research Centre, Oxford Brookes University, Wheatley Campus, Oxford OX33 1HX; jane.hudson@brookes.ac.uk; 01865 485919

Co-authors:

Sue Bloxham, Faculty of Arts, Education & Humanities, University of Cumbria, Lancaster; S.Bloxham@cumbria.ac.uk; Mail to: Green Head Farm, Low Bentham, Lancaster, LA2 7ES; 01524 262037.

Birgit den Outer, ASKe Pedagogy Research Centre, Oxford Brookes University, Wheatley Campus, Oxford OX33 1HX; b.den-outer@brookes.ac.uk; 01865 485736

Margaret Price, ASKe Pedagogy Research Centre, Oxford Brookes University, Wheatley Campus, Oxford OX33 1HX; meprice@brookes.ac.uk; 01865 485944

Abstract: Since their introduction in the 1990s, explicit standards documents have pervaded higher education assessment—success likely linked to their compatibility with constructive alignment and quality assurance regimes. Researchers, however, criticise that such documents are based on a misconception of standards as explicit and absolute, when in fact standards have tacit and contextual qualities that make it impossible to codify them fully. This article considers how practitioners conceive of standards. It identifies the range of concepts of standards, and looks at which were dominant or marginal in 24 external examiners' responses to interview questions about their examining practice. The article identifies a significant gap between the theoretical positions asserted in the research literature and the conceptions held by experienced academics tasked with guaranteeing national standards. It considers implications for quality assurance and reflects on whether the dominance of transparency and accountability discourses leads academics to contort the way they talk about standards.

Keywords: marking, external examiners, situational analysis, assessment standards, learning outcomes

Introduction

'...Modernity has come to mistrust intuition, preferring explicitly articulated assertions; it is uncomfortable with ad hoc practices, opting for systematic procedures, it substitutes *detached objectivity for personal commitment' (Tsoukas 2003, 411).*

The introduction of artefacts into grading practices in the 1990s alongside the rise of quality assurance processes has been a significant intervention in higher education assessment. The range of assessment artefacts, or codifications, has grown to include 'rubrics, criteria-standards matrices, marking guides, scoring schemes, grade descriptors, minimum (threshold) standards, subject or discipline benchmark statements, and graduate attributes' (Sadler 2014, 274-75) as well as learning outcomes. They have been thoroughly integrated into assessment processes and are now seen as 'primary tools for communicating, transferring and sharing "standards knowledge" among learners, academics, accreditation agencies, professional bodies and employers' (Sadler 2014, 275).

Codifications have proliferated to such an extent that in a study published in 2012, Bloxham and Boyd found that markers conceive of standards and criteria as 'almost inextricable' (625). They also observed that markers often talk about standards as internalised. This article unpicks the seeming paradox that standards are both internalised and inextricable from criteria. It extends Bloxham and Boyd's work on practitioners' beliefs about standards by looking closely at how one group of assessors, namely external examiners of taught programmes, characterised standards during interviews about their marking and examining practice. In the UK external examiners are academics selected on the basis of their experience to review the academic standards of programmes at other universities.

Analysing the transcripts of the external examiner interviews, this article asks: in the current context of increasing reliance on codifications, how do external examiners conceive of assessment standards? The article focuses on two dynamics of standards that emerged as salient in our initial analysis: tacitness and contextuality. In other words, it looks at whether external examiners characterise standards as articulable or beyond words; as absolute or situated in the contexts of different programmes and cohorts. It then discusses how external examiners' concepts of standards relate to and differ from the research consensus on the nature of standards and what implications they might have for external examiners' ability to 'ensure that the standard of each award is maintained at the appropriate level' (QAA B7 2011, 8; see also Bloxham and Price 2013).

Tacitness and Codifications

Coordinated efforts to make higher education standards more explicit began in the 1990s, around the time that massification and the Bologna process were reshaping the higher education sector. Outcomes-based assessment was promoted as an alternative to existing assessment, which was seen as 'subjective, anecdotal, even negligent' (Ecclestone 2001, 301). Quality assurance developments focused on transforming the existing 'implicit' model of academic standards, which Stowell characterises as 'centred on essentially undefined assumptions' (2004, 500), through procedures aimed at making standards more explicit (Bloxham and Boyd 2012). Alongside anonymous and blind double marking, codifications were thought to increase the equitability and fairness of assessment (Stowell 2004; Orr 2005). They answered 'calls for more accountability and growing expectations amongst students of "good grades" and guidelines to help achieve them' (Ecclestone 2001, 302).

Codifications fit in with the 'explosion' of audit culture (Power 1994) and the new managerial ethos in UK Higher Education (Hussey and Smith 2002). Hussey and Smith

explain how learning outcomes are designed to be 'observable products' that fit into an auditable system: 'the progress of the student, the suitability of the teaching method and the effectiveness of the teacher, can all be determined objectively: the entire enterprise can be "tracked" and audited, even by someone completely ignorant of the discipline concerned' (2002, 223). The following paragraph from the UK Quality Code for Higher Education offers an example of the dominance of codifications in 2013:

> Individual degree-awarding bodies are responsible for ensuring that UK threshold academic standards are met in their awards by aligning programme learning outcomes with the relevant qualification descriptors in the national frameworks for higher education qualifications. They are also responsible for defining their own academic standards by setting the pass marks and determining the grading/marking schemes and any criteria for classification of qualifications that differentiate between levels of student achievement above and below the threshold academic standards.
> (QAA 2013, 5 emphasis added)

As the word align in the quotation hints, codifications have gained legitimacy from one of the most dominant concepts in assessment of the past 15 years: constructive alignment. Introduced by Biggs in 1999, constructive alignment relies on criterion-referenced assessment and involves aligning assessment criteria with learning outcomes and programme specification. These codifications are seen to create a 'system' in which the learner is 'trapped', and 'finds it difficult to escape without learning what he or she is intended to learn' (Biggs 2003, 2). The codifications system is charged with forcing resistant learners to learn, making teachers interchangeable (Orr 2005) and allowing non-specialists to track student learning (Hussey and Smith 2002). Responsibility for both learning and teaching is placed on the codification system itself; systematic procedures replace personal commitment.

Contextuality and Codifications

A system of aligned codifications relies on an absolute rather than contextual, or relative conception of standards. Criterion-referenced assessment, a key element in constructive alignment, is built on a concept of standard as benchmark, or 'gold' standard independent of context (Bloxham and Boyd 2012). This means that no matter the context, a piece of work with the same quality should receive the same grade. Transparency, equitability and auditability are touted as virtues of criterion-referenced assessment. It is often contrasted favourably with norm-referenced assessment, in which students are measured not against a fixed standard but against each other. In norm-referenced assessment, student performance is ranked and grades are dictated by statistical distribution. In norm-referenced assessment the distribution of grades rather than the relationship between quality and grade remains the same, independent of context. The severance of 'objective' quality and grades means that grades do not flag up contextual factors such as weak cohorts or poor teaching. As such, norm-referenced assessment cannot contribute to the system in which student progress and teaching can be tracked externally (Hussey and Smith 2002).

Criticisms

Researchers have questioned the assumption that codifications have created a fixed 'gold standard'. Firstly, they contest the idea that criterion-referenced assessment does not involve norm referencing (Orr 2008; Bloxham et al. 2011). Norm referencing often takes place informally within supposedly criterion-based assessment; Crisp (2013) found markers sometimes use comparison between student work to fine-tune marks and check their rank order. Secondly, researchers criticise the assumption of even the possibility of fixed standards. Reviewing the literature on standards, Bloxham and Boyd conclude that 'the simplistic and fixed notions of standards as portrayed in public debate deny the necessarily

elusive and dynamic nature of academic standards which are continuously co-created by academic communities' (2012, 617). In other words, standards are contextual but are treated and spoken of as absolute. Sadler explains that the 'qualifiers, modifiers and hedge words in educational standards statements are typically interpreted relatively rather than absolutely' (2014, 281). The terms used in codifications are interpreted—'stretched'—to fit a particular context. Bloxham and Boyd suggest an alternative to the criterion-referenced assessment paradigm, a socio-cultural one in which standards are conceptualised as 'context-dependent, socially situated and interpretive' (2012, 617).

Scholars have also disputed the ability to make standards explicit. Sadler (2014) builds on other research that has revealed fundamental problems with codifications (see Rust et al. 2003; Sadler 2005) when he argues that codifications are 'theoretically incapable of adequately representing standards' (2014). They 'cannot "hold" standards by serving as stable reference points for judging and reporting different levels of student achievement' because 'first, achievement is not a physical variable but a concept which has fuzzy boundaries. Second, the words used to designate amounts are elastic in their interpretation' (275). Hussey and Smith (2002) make a similar argument:

> [Learning outcomes'] alleged explicit clarity, precision, and objectivity are largely spurious. Those academics and teachers who have had to use them have overcome this vacuity either by merely feigning compliance or by implicitly (and perhaps even unconsciously) interpreting them in terms of their existing knowledge and experience (2002, 232).

Codifications allow students and lecturers to collude in what Shay calls 'the myth of objectivity' (2005, 676).

Bloxham et al. (2011), Sadler (2010) and Orr (2010) have all outlined visions for accountability in marking that recognise the tacitness and contextuality of standards. In this vision, what becomes transparent is the fact that standards cannot be completely codified, assessment criteria require interpretation, and expertise needs to be drawn upon to judge quality and maintain standards. Bloxham et al. (2011) and Orr (2008) also suggest recognising the role norm referencing plays as part of acknowledging the contextuality of standards.

While there is a consensus emerging in the research literature about the nature of standards, it is at odds with many established assessment practices and quality assurance processes. External examining is a quality assurance process in which academics negotiate tacit and contextual standards within a regime built on the promotion of transparency and accountability. This article extends the work on the nature of standards reviewed above by looking at how external examiners conceive of standards within this complex field.

Methodology

This paper draws on a Quality Assurance Agency and Higher Education Academy-sponsored research project investigating how external examiners conceive of and apply academic standards. The participants were 24 academics based in 20 different UK universities who externally examine undergraduate programmes in four diverse subjects (chemistry, history, psychology and nursing). They participated in a two-phase interview. In the first hour, they took part in a Kelly's Repertory Grid exercise, which disclosed the qualities they notice in student work, e.g. structure, referencing, engagement with literature. These qualities were then explored in the second hour, an interview and a map-making exercise (modified from Clarke 2005). During the interviews, researchers jotted down on post-it notes any people, experiences, organisations, documents, values or other 'elements' that the examiners

mentioned as influences or explicitly discounted as influences on their standards. In the last 20 minutes of the interview, the examiners used these post-it notes (and any others they cared to add) to create a visual depiction (map) of what they believed to be the provenance of the standards they use. The researchers had each examiner organise these by relative importance around a core and periphery, 'thinking aloud' as they did so. They then asked the external examiners questions about their maps. The entire interview and map-making process was audio-recorded and transcribed. This article focuses on the transcripts of the interview, with attention to the maps created. Outcomes of the KRG exercise are published in Bloxham et al. (2015).

Analysis

In the initial analysis, we approached the data from the interviews and mapping exercises both by participant and by 'element'. We coded all of the interviews using Atlas.ti and summarised the cores of each of the maps, including which elements appeared and how the examiners described them as relating to each other. We also wrote summaries of each of the elements, including information about where and if they appeared on the 24 maps and how they were characterised. Writing these summaries required re-reading and listening to the interview recordings while looking at the social worlds maps to make sure we accurately captured how the examiners characterised each element.

Following Clarke (2005), we next drafted various positional maps. Positional maps look like x-y axis graphs and lay out at the micro level the major positions taken and not taken in the data. They help to articulate a particular issue or controversy around the situation of inquiry, in this case how external examiners conceive of academic standards. We experimented with various concepts on the x and y axes, and in this paper present only one of the maps we created—that which seemed to best illuminate our data in terms of how examiners conceive

of standards. We chose tacitness as an axis because we observed that examiners differed greatly in their commitment to written assessment criteria and other codifications. We chose contextuality as an axis because some examiners were far more concerned with the context of the assignments than others: some put down contextual factors on their maps and some also requested contextual information before being willing to provide a mark during the KRG exercise, while others did not. Juxtaposing contextuality and tacitness on the x and y axes creates a map in which each of the four corner positions represents a distinct 'extreme' position and midway positions that shade into each other (see figure 1).

The positional map represents the 'heterogeneity of positions' (Clarke 2005, 126) available for individuals to take up, including positions that may not be taken up. The unit of analysis is not the individual external examiner; rather, the mapping process identifies 'various social sittings' (Clarke 2005, 126). The analysis recognises that examiners adopt multiple positions in the course of a single interview. It considers which positions (concepts of standards) are outlier, marginal or common, but does not try to locate every examiner in a particular position.

The Tacitness Axis

The y axis on our positional map relates to tacitness. Some terms from the interviews associated with tacitness are sense, emotive sense, internalised sense, judgement, global judgement, belief and instinct. On the other end of the tacitness spectrum is the conception that it is possible to make standards explicit; in other words, it might take work to get the wording perfect, but it is possible to articulate standards in speech or writing. Some words associated with explicitness are objective, transparent and explicit. An explicit concept of standards is often expressed indirectly, through a commitment to codified standards, which are only valid to the extent that standards are explicit.

Tsoukas (2003) argues that the concept of tacit knowledge has been widely misunderstood as a result of Nonaka and Takeuchi's 1995 theory of knowledge conversion, which describes four ways by which tacit knowledge and explicit knowledge are converted or combined. He takes issue with their concept of tacit knowledge as 'knowledge not-yet-converted' and aligns instead with Polanyi's (1962) conception of tacit knowledge as essentially inarticulable. Tsoukas illustrates his understanding with the example of the geographical map, which 'no matter how elaborate […] cannot read itself; it requires the judgement of a skilled reader who will relate the map to the world through both cognitive and sensual means' (Tsoukas 2003, 413 drawing on Polanyi). Although Tsoukas argues that tacit and explicit knowledge 'are not the two ends of a continuum' (2003, 425), we do position tacitness and explicitness at the two extremes of the y axis on the positional map. The axis, however, represents how examiners conceive of standards, rather than our theory of the nature of standards. At one extreme, examiners believe that standards have been made fully explicit, that tacit knowledge has been fully converted to explicit knowledge (taking a Nonaka and Takeuchi view). In the middle, examiners believe that only knowledge's technical part—'that which is possible to articulate in principles, maxims' (Tsoukas 2003, 423)—is captured by explicit standards, while the ineffable is communicated through other means. At the other extreme, examiners believe that standards are so ineffable that codifications are practically useless.

Examiners sometimes speak of standards as 'internalised'. While an examiner could perhaps characterise standards as internalised but articulable, it seems more likely that examiners describe as 'internalised' those aspects of standards that remain after they have taken all external, observable contributions into account. Therefore, tacit and internalised standards are, at a minimum, 'interlinked concepts'—the term Orr (2010) used to describe the relationship between tacitness and connoisseurship.

The Contextuality Axis

The x axis on our positional map relates to contextuality. When standards are viewed as contextual, they are not fixed across cohorts and situations. It is necessary to have contextual information to hand before marking a piece of work. A contextual concept of standards can be seen in the interviews when examiners talk about comparing pieces of student work as part of grading, describe standards as evolving or changing, or indicate that they accept the use of inconsistent standards criteria as appropriate depending on the context. The idea that knowing is inseparable from context is consistent with an understanding of cognition and knowledge as situated (e.g., Lave and Wenger 1991). On the other end of the contextuality spectrum are absolute standards. If standards are conceived of as absolute, it should be possible to pick up any assignment from any cohort and assign it a mark. Language suggesting a single standard—such as discussion about *'the'* standard—suggests a concept of standards as absolute.

[Figure 1: Contextuality/Tacitness Positional Map]

Beginning in the top left corner and moving around the positional map in a counter-clockwise direction, this section looks at each of the hypothetical positions presented by the positional map above, both the extreme positions located in the four corners of the map as well as the in-between positions. It considers what each intersection of tacitness and contextuality might look like in practice—whether it might be associated with internalised standards or codifications, and how. The map is then used as a framework for thinking about the data, including an analysis of which positions were taken up by the external examiners within the interviews. The analysis refers to the examiners by number, from Examiner 01 through Examiner 24 (more details about the examiners can be found in Bloxham et al. 2015).

Absolute/Tacit

If standards are both absolute and tacit, they are stable across contexts and cohorts but cannot be articulated. When an academic says he cannot explain exactly how, but 'I know a 2:1 when I see it', he is speaking from a concept of standards as tacit and absolute (for more perspectives on this claim see Ecclestone's 2001 article with this title). If this sounds old fashioned, it is probably because 'relatively fixed but undefined standards' have been the 'only model of academic standards prevalent in higher education until fairly recently' (Stowell 2004, 500). Because standards are completely tacit, codifications are of no value in the extreme 'absolute/tacit' position. Lower down on the tacitness axis is the 'absolute/somewhat tacit' position. Here, codifications may have some utility, perhaps as aides-memoires for technical aspects of standards that can be made explicit.

External examiners' concepts of standards as absolute/tacit in the interviews

The absolute and tacit concept of standards, so dominant two decades ago, was still drawn upon relatively often by external examiners. This concept was seen most clearly when external examiners identified moderation processes (Ex03, 06, 07) or relationships with mentors (Ex01, 19, 20) as having been influential on their standards early in their careers. Crucially, they no longer consider these processes to influence their standards. When they engage in moderation now, they do so to help others establish their standards. They described their mentoring as a training role in which they impart knowledge rather than as an opportunity for two-way sharing or calibration. Their perception that sharing processes and relationships are useful in establishing standards indicates that they conceive of standards as at least somewhat tacit. Once these standards are established, however, they remain fixed and engaging in processes to adapt them to different contexts or to keep them in line with peers is unnecessary. Many external examiners talk about how they used to be able to change marks

on individual assignments as part of their examining practice. Such a role would be consistent with an absolute/tacit concept of standards.

Examiners sometimes spoke about standards in a way that indicated they conceived of them as absolute and having some qualities that could be made explicit. Examiners 04 and 08 commented that national subject benchmarks had become less important as they had grown in experience, suggesting that they felt they had internalised these. Examiner 19 reported that unlike some of her younger colleagues she does not mark with the criteria in front of her, but she does pull them out when she cannot place a specific piece of work or when she needs support in explaining a mark to a student. Other examiners also described codifications as tools they could use to jog their memories, for example when they need to remind themselves of the difference between levels (Ex08) or 'not to get beguiled by the surface features' (Ex03).

Absolute/Explicit

The 'absolute/explicit' position in the bottom-left corner of the positional map holds that standards can be precisely articulated. Thus, in the 'absolute/explicit' concept of standards it would be possible to write standards documents for a given level that were relevant across the country and across cohorts. For this reason, this position is compatible with a commitment to national or discipline-wide level descriptors and a belief that all standards documents in a given field align with one another. Absolute and explicit standards could theoretically also be shared verbally in a straightforward and accurate way. This concept of standards is frequently found in quality assurance documents (for example, see the UK Quality Code quotation in the introduction of this article). Scholars of higher education standards have argued that explicitness and absoluteness are incompatible: 'the meaning of evaluative terms used to

specify the quality of knowledge, understanding or analysis are always relative to a context and so cannot be used to specify absolutes' (Hussey and Smith 2002, 229).

External examiners' concepts of standards as absolute/explicit in the interviews

Examiner 13 illustrates the belief in the explicitness of standards associated with the 'absolute/explicit' position when he says that national benchmarks can be defined 'very precisely'.

> I always do my external examination stints considering all those benchmark statements and they define very precisely how a level 4 student should be and what are the needs, the features of level 5 students and level 6. (Ex13)

Examiner 13 describes his role as bringing together all the codifications to make sure he achieves every aspect of 'the standard.' He describes external examining as being a 'guru of all the handbooks'. Most examiners, however, said that they rarely referred to national benchmarks in their examining practice. The 'absolute/explicit' concept of standards was expressed most often in response to questions about the role of national benchmarks in the interviewees' examining practice; external examiners explained that they trusted that local standards artefacts had already been aligned to national reference points. Examiner 01 commented "I think they [national reference points] fed into things like this [institutional] guide, I personally rarely refer to them" (Ex 01). Examiner 03 put HEFCE benchmarking on the periphery of his map because 'it influences the assessment outcomes, the grade descriptors and the level descriptors [...but] we don't think of it directly.' External examiners did not see their role as checking that national and local criteria were aligned. Only examiner 04 hinted that the role of the examiner would 'ideally' include picking up and ameliorating bad criteria so that she could go on 'to objectively apply the standards and to make sure that everything is equal.'

Contextual/Explicit

Standards conceived of as explicit and contextual (or somewhat contextual) can be fully articulated but are shaped by and situated within their specific contexts. This concept allows for the coexistence of multiple standards; codifications are not necessarily all aligned. The 'contextual/explicit' concept of standards underlies the view that an examiner's role is to know which codified standards are appropriate to apply in a given situation and to do so, leaving his or her personal judgements out of the process.

External examiners' concepts of standards as contextual/explicit in the interviews

There were many instances in the interviews of examiners characterising standards as 'contextual/explicit' or 'somewhat contextual/explicit'. This was often expressed as a strong commitment to relying on local assessment criteria. The quotation below illustrates how the contextual/explicit position was expressed:

> I am a firm believer that when I am externally examining I am working to one set of marking criteria and so on, and if that's what you have said to the students that's what they will be judged against, that's what you've got to use. When I'm here, as I say, it's against ours, when I am there it is against theirs. (Ex11)

By completely discounting the role of her own judgement this examiner suggests that she conceives of standards as fully captured by marking criteria. Her promise to work to the local marking criteria exclusively suggests that she conceives of standards as contextual. In this quotation she does not indicate a belief that all codifications should or do align.

Contextual/Tacit

In a 'contextual/tacit' concept of standards, standards are elusive. They cannot be articulated precisely and they emerge in specific contexts. Researchers have argued that this concept of

standards is most accurate (e.g., Hussey and Smith 2002; Bloxham and Boyd 2012; Sadler 2014). In this concept, one way to arrive at standards is to negotiate them in local contexts on a constant or at least regular basis. Sadler recommends sharing standards by joint participation in evaluative activity (Sadler 2005).

Related conceptions of standards radiate out from the upper-right hand corner 'contextual/tacit' position. In the 'contextual/somewhat tacit' position, documents are a blunt but useful tool for some aspects of standards within these negotiation processes. The 'somewhat contextual/somewhat tacit' position in the middle of the map could represent the view that standards documents are a general guide to standards and their interpretation should be influenced by the context and cohort.

External examiners' concepts of standards as contextual/tacit in the interviews

A firmly expressed 'contextual/tacit' position was absent in the data. Examiners 01 and 19, however, both put 'actively comparing student work' near the core of their maps and Examiner 01 said that her standards changed over time influenced by 'what she saw'. The 'somewhat contextual/somewhat tacit' position could be seen when examiners 18 and 19 mentioned being influenced by 'a fresh pair of eyes' during moderation, but even such cautiously flexible comments were rare. Examiner 10 described the external examiner process itself as a calibration activity, characterising it as helping 'keep your standards tip-top' and giving 'a measure of if you're still on top of your game'. In general, however, examiners described moderation and mentoring as tools for establishing internalised standards for new markers, not as sites and moments from which standards emerge and are agreed.

Examiner 24 expressed a concept of standards as 'somewhat contextual/somewhat tacit' in response to a question about the importance of national reference points: 'I suppose I'd be aware of them, I think I've got some of them somewhere … so I think again it's probably part

of this idea that there is this internalised sense but then that's balanced by external things and I think that would be one of the external things.' Examiner 21 attributed his ability to be a 'good, consistent marker' to mentoring and training in marking and marking criteria, including sampling work and having 'quite a lot of complicated and in-depth and robust conversations about grading.' He sees marking criteria as useful but relies on mentoring to give him a 'sense' of how to mark with them and on robust conversations to be able to mark consistently. Examiner 21 believes there is a somewhat fixed standard but processes are necessary to support consistent interpretation of assessment criteria.

Discussion: The Dominant Positions

The external examiners interviewed frequently expressed beliefs that relying entirely on explicit and traceable processes for judging student work is most fair, objective and transparent. Examiner 08 is committed to criteria because she believes it increases transparency and helps students understand how to do better.

> I think for me it's important that they also understand the taxonomy and the expectations of how they're going to be marked, because I'm a great believer in clear and transparent criteria for students to use based on how they can achieve, so how can they use their feedback to get onto the next level of a taxonomy, for example. So they're not perpetually sitting in the 40s or 50s without understanding what does it mean to begin to critique or begin to critically analyse. (Ex08)

Examiner 08 imagines a rational process whereby feedback, taxonomy and criteria are part of a system that is transparent to both students and teachers. It is worth noting that the quotation from examiner 08 points to one of problems with explicit standards documents outlined in the literature: that terms like 'critique' and 'critically analyse' are not in fact self-evident—they need to be 'understood'. As Hussey and Smith argue, words like these 'give the impression of

precision only because we unconsciously interpret them against a prior understanding of what is required. In brief, they are parasitic on the very knowledge and understanding that they are supposed to be explicating' (2002, 225).

While external examiners evoked various, sometimes contradictory forms of fairness in their interviews (e.g. providing the public with accurate standards, taking into account student backgrounds, equitability of time spent on scripts), the form they expressed most concern for was fairness to the students in terms of marking to the criteria the students had been given. As a result, they were committed to marking to local codifications. External examiners characterised standards 'contextual/explicit' more often, forcefully and spontaneously than they characterised them as 'absolute/explicit', despite the fact that the latter position is often assumed in quality assurance processes. This distinction between a conception of explicit standards as contextual rather than absolute might not have surfaced had we interviewed general markers, because the interviewees' beliefs that standards are contextual emerged as they talked about marking (and using codifications) across programmes and institutions. The examiners were dedicated to local assessment criteria, but did not see them as necessarily aligned with one another across programmes or institutions.

Examiner 04 conceives of her role as examiner as checking the work of local markers who may have wrongly relied on implicit, uncodified standards.

> It comes down to the objectivity for me because the marking team will know their criteria, but typically when they mark they don't mark directly against the criteria, they'll have a sense of what they're looking for and they'll go in and look for that. I want to know that actually it does stand up against the criteria that they have dictated themselves. (Ex04, emphasis added)

Examiner 04 contrasts her own 'objectivity,' which she associates with marking to the criteria, with the marking team relying on their 'sense of what they're looking for'—an approach she clearly believes undermines fair assessment. Unlike reliance on 'sense', reliance on explicit standards is seen as 'untainted by values, culture or power' (Bloxham and Boyd 2012, 617).

In contrast, examiners were reluctant to say that personal, informal experience that they had internalised shaped their standards. Looking at positions taken and not taken in the data—rather than associating each respondent with a single concept of standards—highlighted that the interviewees actually drew upon the traditional and maligned 'I know a 2:1 when I see it' concept of standards relatively frequently, even as they disavowed it at other moments in their interviews. When examiner 24 linked her own university experiences to her sense of standards, she said apologetically, '...which I suppose is where a lot of it probably comes from, *if I'm honest*.' Examiners often apologised or adopted a confessional tone when they broke from the discourse of transparency and objective criteria, suggesting they believed that such language constitutes the 'correct' way to talk about standards. Despite these reservations, the concept of internalised tacit and fixed standards was common, and often appeared where examiners talked themselves out of their commitment to explicit standards.

Within the interviews, there were several remarkable instances of external examiners retracting their strongest assertions of dedication to codifications. Examiner 04 stated that she would like to think her commitment to objectivity was such that if she were 'given something that had appalling criteria that [she] would rigorously apply those appalling criteria to the pieces of work that [she] was looking at, because that is what's fair actually.' Shortly afterward in the interview, however, examiner 04 said about being ruthlessly objective: 'now I'm thinking about it I feel… it's a bit like a Nazi officer saying, "Well my job is to do this,

therefore I should do my job.'" Examiner 02 was also very committed to using assessment criteria, explaining 'there's almost like a contractual agreement with students' when that is how they've been told they will be marked. She confessed, however, 'there will come a point at which my training will say to me, no, your beliefs are right, go with your instincts and say you've got to.' She was committed to local standards documents as her 'starting point', but she seems actually to conceive of standards as internalised—fixed and tacit—and felt this deeper standard 'instinct' would override explicit standards where they were in contradiction. Examiner 04 also acknowledged that her subjective interpretation has a role to play in applying explicit criteria. While Examiners 04 and 02 both believed that they 'should' rely exclusively on codifications, they both also felt that other, personally held values or internalised standards would prevent them from deferring to assessment criteria in every instance. In short, they would speak the language of the 'contextual/explicit' position but only insofar as the assessment criteria were loosely in line with their own absolute standards.

Despite encountering problems with talking about standards as if they can be made completely explicit, at moments during many of the interviews external examiners protested that they do not bring their own standards to the table and that they rely instead on marking criteria. Recall the statement, 'When I'm here, as I say, it's against ours, when I am there it is against theirs' (Ex11). The external examiners distanced themselves from uncalibrated internalised standards, concurring, it seems, that 'there is a fine balance between a genuine ability to recognize quality of work apparently intuitively, and erratic interpretation' (Ecclestone 2001, 305). Although they sometimes backtracked on their statements, the interviewees asserted the 'contextual/explicit' concept of standards more confidently than any other of the positions laid out by the positional map.

Conclusion

External examiners' conceptions of standards have quality assurance implications. An external examiner operating within either the 'absolute/tacit' or the 'contextual/explicit' concept of standards would theoretically be unable to contribute to comparable standards nationally. In the former concept, there is no mechanism to guarantee consistency between examiners (Bloxham and Price 2013). In the latter concept, external examiners cannot contribute to national standards because each examiner merely makes sure that local markers follow the criteria that those markers themselves set down. The system of aligned codifications is a closed loop and the external examiner merely ensures it is following its own rules. Comparable national standards would be possible within an 'absolute/explicit' concept of standards—in such a world, an interchangeable 'external examiner' might merely observe that the local processes of the transparent national system were functioning as designed. Critically, of course, this techno-rationalist dream leaves aside the inconvenient problem of the fuzziness, elasticity and situatedness of standards.

The concept of standards as tacit and contextual, requiring constant, situated negotiation was marginal in the interviews despite being well supported in the research literature (e.g. Bloxham and Boyd 2012). The gap between practitioner conceptions and theoretical positions suggests that the language of tacit and contextual standards has not trickled down from the research literature into the broader discourse of higher education assessment, even to those experienced academics charged with guaranteeing national standards. A concept of standards as fluid and unobservable may be untenable—and thus unthinkable—within an audit culture. Indeed, there is no metaphor for socially situated tacit standards; words such as internal, instinct and sense are normally associated with individuals rather than groups. Shay (2005) borrows the term 'intersubjectivity' to describe the socially situated ability of

communities to interpret student performance (668), but there is no ready vernacular term for such an idea.

In contrast, explicit standards documents fit in with audit processes and constructive alignment, are in line with the dominant values of transparency and accountability, and are seen to contribute to equitability of assessment. All of this support may lead some external examiners to feel that relying entirely on explicit and traceable processes for judging student work is the most 'correct' way to mark and external examine.

Talking about standards as if they are completely explicit required the conceptual acrobatics and the backtracking that we saw in the interviews. This difficulty may result from having 'to negotiate a techno-rational view of standards as portrayed through QA bodies and institutional quality procedures and the private and tacit experience of embedding standards in their marking' (Bloxham and Boyd 2012, 621). The dominance of transparency and accountability discourses presents a challenge to external examiners' ability to reflect on the nature of standards and how they could be shared. Scholars of assessment standards have, however, already outlined how the values of transparency and accountability could be preserved while also giving up the charade of relying entirely on explicit standards documents (Bloxham et al. 2011; Sadler 2010; Orr 2010). For examiners and their communities, a new accountability would mean ensuring that examiners can apply tacit standards consistently, by engaging and being seen to engage in processes within their disciplines that are focused not on procedural moderation but rather on sharing understandings of standards and developing discipline norms.

Recognition of contextuality and tacitness of academic standards underpins external examiners' ability to contribute to comparable standards across institutions. Many of the

contradictions the interviewees struggled with have been resolved in the research literature, but such literature has not shaped the way they talk about standards or their role, nor relieved them of the stress that comes from forcing standards, which are by nature tacit and contextual, to make sense within assessment systems and quality assurance processes valued for their transparency and objectivity.

References

Biggs, J. 2003. "Aligning Teaching for Constructing Learning." HEA. https://www. heacademy.ac.uk/sites/default/files/resources/id477_aligning_teaching_for_constructing_ learning.pdf

Bloxham, S. 2012. "'You Can See the Quality in Front of Your Eyes': Grounding Academic Standards between Rationality and Interpretation." Quality in Higher Education 18 (2): 185– 204.

Bloxham, S., and P. Boyd. 2012. "Accountability in Grading Student Work: Securing Academic Standards in a Twenty-first Century Accountability Context." British Educational Research Journal 38 (4): 615–634.

Bloxham, S., P. Boyd and S. Orr. 2011. "Mark my Words: The Role of Assessment Criteria in UK Higher Education Grading Practices." Studies in Higher Education 36 (3): 655–670.

Bloxham, S., B. den Outer, J. Hudson and M. Price. 2015. "Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria." Assessment & Evaluation in Higher Education. doi: 10.1080/02602938.2015.1024607

Bloxham, S. and M. Price. 2013. "External Examining: Fit for Purpose?" Studies in Higher Education. doi: 10.1080/03075079.2013.823931

Clarke, A. E. 2005. Situational Analysis: Grounded Theory after the Postmodern Turn. Thousand Oaks, CA: Sage.

Crisp, V. 2013. "Criteria, Comparison and Past Experiences: How Do Teachers Make Judgements when Marking Coursework?" Assessment in Education: Principles, Policy and Practice 20 (1): 127–144.

Ecclestone, K. 2001. "'I know a 2:1 when I see it': Understanding Criteria for Degree Classifications in Franchised University Programmes." Journal of Further and Higher Education 25 (3): 301–313.

Hussey, T. and P. Smith. 2002. "The Trouble with Learning Outcomes." Active Learning in Higher Education 3 (3): 220–233.

Lave J. and E. Wenger. 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge: University of Cambridge Press.

Nonaka, I. and H. Takeuchi. 1995. The Knowledge-Creating Company. New York: Oxford University Press.

Orr, S. 2010. "'We kind of try to merge our own experience with the objectivity of the criteria': The Role of Connoisseurship and Tacit Practice in Undergraduate Fine Art Assessment." Art, Design and Communication in Higher Education 9 (1): 5–19.

Orr, S. 2008. "Real or Imagined?: The Shift from Norm Referencing to Criterion Referencing in Higher Education." In Balancing Dilemmas in Assessment and Learning in Contemporary Education, edited by A. Havnes and L. McDowell, 133–144. New York: Routledge.

Orr, S. and M. Blythman. 2005. "Transparent Opacity: Assessment in the Inclusive Academy." In Improving Student Learning: Diversity and Inclusivity, Proceedings of the 2004 12th International Symposium, edited by C. Rust, 175–187. Oxford: OCSLD.

Polanyi, M. 1962. Personal Knowledge. Chicago: The University of Chicago Press.

Power, M. 1996. The Audit Explosion. London: White Dove Press.

QAA. 2013. "Setting and Maintaining Academic Standards." Part A in UK Quality Code for Higher Education. Gloucester: QAA.

QAA. 2011. "External Examining." Chap. B7 in UK Quality Code for Higher Education. Gloucester: QAA.

Rust, C., M. Price and B. O'Donovan. 2003. "Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes." Assessment & Evaluation in Higher Education 28 (2): 147–164.

Sadler, D. R. 2014. "The Futility of Attempting to Codify Academic Achievement Standards." Higher Education 67 (3): 273–288.

Sadler, D. R. 2010. "Assuring Academic Achievement Standards at Griffith University (2nd Ed." Griffith University. www.griffith.edu.au/__data/assets/pdf_file/0010/375337/Assure AcadStndrds-Consultation-Paper.pdf

Sadler, D. R. 2005. "Interpretations of Criteria-Based Assessment and Grading in Higher Education." Assessment & Evaluation in Higher Education 30 (2): 175–194

Shay, S. 2005. "The Assessment of Complex Tasks: A Double Reading." Studies in Higher Education 30 (6): 663–679.

Stowell, M. 2004. "Equity, Justice and Standards: Assessment Decision Making in Higher Education." Assessment & Evaluation in Higher Education 29 (4): 495–510.

Tsoukas, H. 2003. "Do We Really Understand Tacit Knowledge?" In Handbook of Organizational Learning and Knowledge Management, edited by M. Easterby-Smith and M. Lyles, 411–427. Cambridge, MA: Blackwell.

Figure 1:
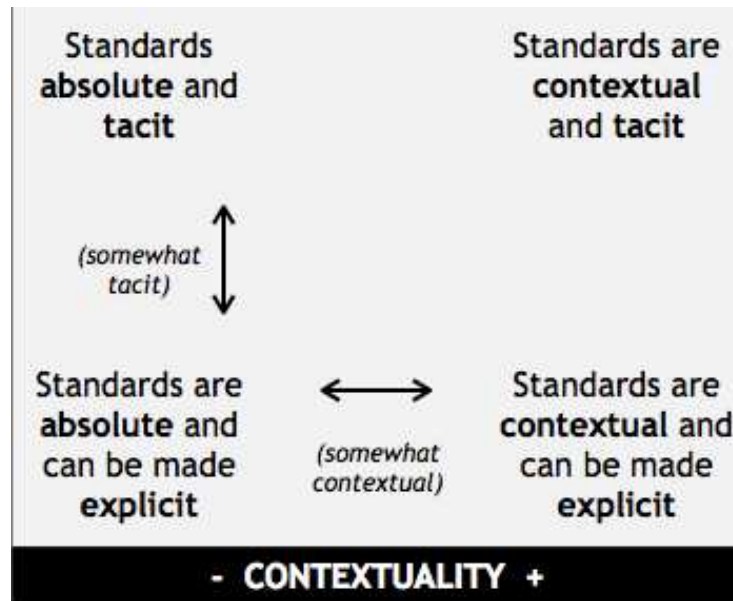
Figure list

Figure 1: Contextuality/Tacitness Positional Map