'Pigs are not fattened by being weighed' – so why assess clinic- and can we defend our methods?

Carol Boothby¹

Northumbria University, UK

For those clinics that assess their students, there can be a panoply of issues to consider. The nature of clinic means that the experience of students is non-standardised, not least in terms of workload. Is it appropriate to assess such an experience? How can clinical teachers be sure that their assessment methods are valid and reliable?

WHY ASSESS IN CLINIC?

Perhaps because teaching and participating in a clinical experience can take such a wide variety of forms, the approaches to assessment have been similarly diverse. Many law schools have students involved in a range of pro bono activities, the majority of which will not be assessed. According to the LawWorks Law School Pro Bono and Clinic Report 2014,² of those law schools that responded to the survey, 96% do pro bono work. This report suggests that (in the UK at least) clinics are increasingly becoming assessed as a credit bearing part of the curriculum and whereas previously only 10% of law schools in 2010 assessed student performance,

¹ Carol Boothby is Director of the Student Law Office at Northumbria University

today this total is around 25% - relatively low, but apparently increasing, perhaps as clinics move from extra to intra curricular.³

Views diverge on the value of assessing clinic, as well as how to do so. In terms of the views of students, recent work by Combe indicates a minority (40%) responded negatively to the question “would you feel comfortable being assessed on law clinic work?”, suggesting that the majority were ‘either perfectly happy or indifferent to the prospect of assessment’.⁴ Brustin and Chavkin also found that the “overwhelming majority of the students believed that clinical courses should be available on a graded basis”, one student commenting that ‘grading permits rewarding those who make greater effort and excel…’.⁵ Other writers do not challenge assessment per se, but challenge the idea of grading. Rice argues that;

‘Grading undermines the collaborative role of the clinical teacher. This is not a journey where we arm students with a map and compass drop them in the wilderness, and give a prize to the first one home. This is a journey we travel with them, clearing the path ahead, holding back to

⁴ Combe, supra n. 2, at p. 292
let them go ahead, offering them a steadying hand, coaxing them on narrow bridges over deep ravines, exhorting them to climb steep hillsides… grading distracts us from our teaching’.  

Hyams disagrees, seeing the reluctance to grade as ‘an evasion of our duty to our students’ and Levine sees pass /fail as not providing enough feedback to enable improvement.  

Nelson and Murray, reviewing the move to the use of grade descriptors at the clinic here at Northumbria Student Law Office, also challenge the case for pass/fail in clinic, arguing that grading recognises the efforts displayed by students and it motivates them to achieve.  

Perhaps the idea that grading distracts from teaching is more likely where the assessment is summative in nature. Where supervisors are providing ongoing formative feedback, and where the method of assessment is fully aligned with the clinical work, assessment can drive learning. From the clinical supervisor’s point of view, one reason for assessing and grading could be that it isn’t enough to simply get students to a ‘pass’ level - we are wanting to help students to move along a continuum towards being ‘practice ready’ - and perhaps for them also to have some awareness of how near or far they are from that. Stuckey’s definitive work, The Best

---


Practices Report, suggests that assessment methods may have more influence on how and what students learn than any other factor.\textsuperscript{10} The benefit of assessment in providing information to students is touched on by Foxhoven;

‘Assessment is a powerful tool because law students uniformly desire to be prepared to become competent lawyers, but, being novices in the legal profession, they are unable to identify core competencies themselves.’\textsuperscript{11}

Like a runner who checks their times and receives advice from their coach in order to improve their performance, students in clinic can (depending on the nature of the clinical experience) use formative assessment feedback to improve their performance. As Brown and Knight argue, ‘far from it being the case that you’ll not fatten a pig by weighing it… the science of weighing is necessary for the art of development’.\textsuperscript{12}

Assessment provides information about student learning – but a stronger claim (according to Brown and Knight) is that assessment shapes the curriculum; ‘Assessment defines what students regard as important ’.\textsuperscript{13}

Coffield et al, in a comprehensive and critical examination of learning styles, refer to the work of Desmedt in finding that, ‘because of the curriculum, students are not


\textsuperscript{11} J. R. Foxhoven, “Beyond Grading: Assessing Student Readiness to Practice Law” (2009) 16 Clinical Law Review 335, p. 344

\textsuperscript{12} S. Brown and P. Knight, “Assessing Learners in Higher Education” (London, Kogan Page, 1994) p. 46

\textsuperscript{13} Ibid., at p. 12.
interested in learning, but in assessment.’ 14 This may seem a depressing indictment of students, but surely it is not specific to students, but simply of human nature; if the way in which a race is run is judged on time, then no matter how much we may exhort a particular running style, unless this actually contributes to the goal of ‘best time’, it is likely to be discarded or ignored. For many clinicians in law clinics, particularly those driven by a social justice motive, (as Rice is) there is a risk that, unless the assessment and any grading leads to and measures progress in social justice terms, these aspects are merely a distraction.

SO WHAT IS ‘GOOD’ ASSESSMENT?

Arguably, a crucial factor underpinning all the support for assessing is how useful the assessment actually is in driving learning. What do we mean by ‘useful’? Taking forward the point about concerns over validity of assessment, this can be a perplexing area. One field that legal clinicians can (and have) drawn from is the medical profession. The use of problem based learning in the teaching of law has been derived in this way, as was the use of standardised clients and the training used in medicine continues to provide a rich seam of expertise.

Those assessing medical students have puzzled over many of the same issues as legal clinicians. In particular, the work of Van der Vleuten, an academic in the field

not of medicine, but of education, led to him becoming the “accidental hero” ¹⁵ of medical education, who wanted to discover “promising ways to advance and to prevent repetition of the mistakes of the past in the future”, ¹⁶ moving away from high stakes assessment.

In summary, Van der Vleuten uses a conceptual model for confirming the ‘utility’ ¹⁷ - simply put, the ‘usefulness’ - of an assessment method by using a mathematical model incorporating key aspects such as validity, educational impact, and acceptability. This model can help us to examine what good assessment in clinic might look like, and this is a process which has been started at Northumbria University’s in-house law clinic, the Student Law Office.

**ASSESSMENT AT NORTHUMBRIA STUDENT LAW OFFICE**

Can the reflective and experiential elements of CLE be codified into assessment rubrics that provide guidance to students (and staff) without reducing their depth and complexity? At Northumbria, the law clinic moved in 2007-8 from criterion-referenced, outcome-focused assessment to the use of 10 grade descriptors, including a range of skills and attributes from oral communication, written communication, to key skills such as a student’s ability to demonstrate autonomous

---

learning. 18 There is no explicit ‘justice’ agenda within the clinic curriculum, although it is implicit in many of the activities and experiences. So how is the clinical training and assessment of students linked to the wider discourse of what a lawyer is, and can be? Through the development of reflection and reflective / reflexive practices, students have the opportunity to consider their own development, (see appendix A for assessment matrix). In addition to carrying out casework under supervision, and being assessed on these through the grade descriptors, students are currently required to produce two reflective pieces, 2000 words each, one on the development of their skills, and a second drawn from a list of topics such as ‘Law in Action’, ‘Clinic and my Career’, ‘Justice and Ethics’ and ‘Clinic and Legal Education’. These reflections are submitted with the portfolio evidencing their casework, at the end of the module.

If challenged, how could we defend our use of our current form of assessment? Attacks can come from either end of the spectrum- those who see the social justice mission as too important for things like assessment 19 and grading 20 to get in the way, and others who worry that the experience of clinic is too non-standard and that this variety of experience needs to be narrowed into a check list of activities. The writer experienced such a challenge from an external examiner, who questioned the variability of the clinic experience, and the lack of control staff or students have over

---

18 Murray, supra n.7, pp. 48-60.
19 Nicolson, supra n.2.
20 Rice, supra, n.5.
this, which provided much food for thought, reflection, and a useful opportunity to critique and justify our existing methods. It brought a realisation that it may not be enough to rely on the mantra of clinic being so good that challenging assessment validity is a heresy. On the other hand, it cannot be that simply because clinic is a non-standard experience, (arguably one of the reasons students engage with it) and because this makes assessment difficult, that we give up either on clinic within the curriculum or assessment of it. Being able to deconstruct and critique clinical methods, including assessment tools, should help to understand our clinical teaching more deeply- and perhaps also to see it from the students’ point of view, in terms of alignment and authenticity.

So if non standardisation is one purported challenge to the validity of assessment in clinic, what are the other potential components of validity?

Van der Vleuten’s work based on the ‘utility model’ gives a framework within which to carry out a methodical examination of our use of assessment.

THE UTILITY MODEL

Van der Vleuten uses the idea of ‘utility’ as a conceptual model, whereby criteria are multiplied together to produce a utility index. Those criteria can include;

1. Validity (does an assessment instrument measure what it purports to?)
2. Reliability (can scores for an assessment be reproduced )
3. Educational impact – the impact of assessment on learning
4. Acceptability to stakeholders/Cost – in terms of resources
He makes 2 further points; that

- Selecting assessment methods involves context-dependent compromises
- Assessment is not a measurement problem but an integrated design problem\(^{21}\), made up of educational, implementation and resource aspects.

What was already known was that the usefulness of assessment depends on compromise between various quality parameters, but what Van der Vleuten highlights in his later work is that:

1. ANY method of assessment may have ‘utility’, depending on use
2. We need more methods using qualitative information relying on professional judgement, the latter being highly valuable.
3. Assessment is an ‘educational design problem’\(^{22}\) that needs a holistic approach.

In terms of reliability, here lies the importance of sampling, by which Van der Vleuten appears to mean that because competence is highly dependent on context or content, we need to use a large sample across the content of the subject to be tested, particularly if there are other potential effects on reliability, such as, in the case of clinic, clients, and single supervisor. This has relevance for assessment at Northumbria SLO, where it could be argued that, through the use of a wide range of grade descriptors, and ongoing assessment, we compensate to some extent for the ‘single supervisor’ aspect- but is that enough? The Northumbria SLO clinic assessment includes a thorough moderation process, where a sample of each supervisor’s marking is examined by a different supervisor. But we do not directly

\(^{21}\) Van der Vleuten, *supra* n.17, p.310.
\(^{22}\) Van der Vleuten, *supra* n.18, at p.314.
involve more than one supervisor in the clinic assessment. Further, there is no real link between the practical work, the grade descriptors and the 2 reflective pieces, save that these pieces purport to be a reflection on the clinical experience. In reality, students treat these as an end point assessment, and for many it seems to take until the end of the clinic module for them to grasp what is required. Therefore the use of reflections as part of the clinical assessment is currently being re-examined, and Van der Vleuten’s framework has provided a useful structure.

**ASSESSING REFLECTIVE PRACTICE**

There is widespread use of reflections as a key part of clinic assessment; some clinics incorporate a presentation as well as written work, but no assessment of casework carried out by students. Others are all written but include formative pieces such as a ‘critical incident’ report. All appear to embrace the concept of reflection with gusto, although there have been critiques of the use of reflections. In looking to apply Van der Vleuten’s work on assessing competence, the area of student reflection is one which has been of concern to clinic staff. In 2013, the writer introduced the reflections matrix (Appendix B) to Northumbria SLO assessment,

---


providing rubrics to help clinic supervisors in assessing, and feedback from these assessors was generally positive, but other aspects of the way in which reflection was assessed remain the subject of concern, for example, the structure of the reflections as being essay-style pieces, submitted at the end of the module.

Testing our assessment in clinic against the utility model using grid structured questions provides a structure for discussion. As an overview, the grid below can help to summarise Van der Vleuten’s approach, and enable a critique of current or proposed assessment practice.

<table>
<thead>
<tr>
<th>Element of trustworthiness</th>
<th>Criteria</th>
<th>To what extent is this achieved through current assessment strategy in the Northumbria clinic?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credibility (internal validity)</td>
<td>Aligned with stage of competency (Miller’s triangle)</td>
<td>Complex tasks/requiring mastery of skills, similar to legal practice</td>
</tr>
<tr>
<td></td>
<td>Authentic integration of competencies at each stage</td>
<td>Good integration of legal skills</td>
</tr>
<tr>
<td></td>
<td>Structural coherence within the programme</td>
<td>Grade descriptors align to the skills required for clinical casework</td>
</tr>
<tr>
<td></td>
<td>Prolonged engagement, triangulation and member checking</td>
<td>Good training of assessors (clinic supervisors)</td>
</tr>
<tr>
<td>Transferability (generalizability)</td>
<td>Time sampling</td>
<td>Judgement based on broad sample of data points, repetition of tasks</td>
</tr>
<tr>
<td></td>
<td>Thick description</td>
<td>Assessors justify decisions in detail</td>
</tr>
</tbody>
</table>
For supervisors less familiar with the terminology used by Van der Vleuten, a more user friendly approach asks the question; ‘What would failure to meet/meeting/exemplary practice in relation to this criterion look like?’.

A pilot group using this table plus a brief explanation of Van der Vleuten’s work were able to engage with a valuable critique of our current assessment of reflections. Points raised in relation to the current clinic assessment at the Student Law Office were:

**Competency**- we would expect students to be able to reflect at a reasonably sophisticated level – but have we provided sufficient previous experience and support to raise their reflective skills to the level of study they were at, which is Masters level (level 7)?

**Integration of competencies** - the use of end-point essay-style pieces for assessment of reflection separates reflective practice from the ongoing development and mastery of complex legal skills, so that learning and development of competency in reflection
is not perceived by students in the same way as their development of those legal skills. This is reinforced by the contrast between the high level of formative feedback provided for practical casework, and the limited opportunities built in to the assessment for the purposes of reflection.

**Structural coherence within the programme** - The reflections matrix sets out the way in which the written piece will be marked, but this does not link to or facilitate an ongoing reflective process and perhaps fails to assess authentic growth in reflective skills. At a programme level, it could be argued that there is little prior preparation for the development of reflective skills.

**Prolonged engagement, triangulation and member checking** – the current perception of the reflective pieces as ‘end point’ led them to be summative in nature. In reality, students can prepare them during the year, but the only point at which they have the opportunity to gain supervisor feedback is at the mid year appraisal, when students submit a one-page draft. There is little triangulation, in the sense that the reflections are freestanding pieces of writing. The use of a different format such as presentations might provide an opportunity to engage with students directly and assess the level of true understanding and genuine reflection.

**CONCLUSIONS**

The good news for clinicians is that, as Van der Vleuten says, there is ‘no need to banish from our toolbox assessment instruments that are rather more subjective and not perfectly standardised, on condition that we use them sensibly and expertly. We
can move assessment back to the real world of the workplace as a result of the
development of the less standardised but nevertheless reliable methods of practice
based assessment’.26

Authenticity is valued, as is the role of professional judgment by those assessing.
Tasks should be treated in a holistic rather than reductionist way. We need to
‘construct an overall judgment by triangulating information across these sources’ 27—
perhaps something analogous to the way in which judicial judgements are reached.
A thoughtful and informed approach to assessment in authentic learning
environments such as law clinics should enable this assessment process to be both
informative in terms of student development and reliable as a measure of
achievement.

26 Van der Vleuten, supra n. 18, p. 312.
27 Van der Vleuten, supra n. 18, p.313.
Appendix A: Grade descriptor for Student Law Office (The assessment criteria are equally weighted.)

<table>
<thead>
<tr>
<th>Assessment Criteria</th>
<th>Grade descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F(below 50) 2:2 (50-59) 2:1(60-69) 1st (70-79) +1st (80+)</td>
</tr>
<tr>
<td>Autonomy and efficiency</td>
<td>Poor initiative shown; routinely relies on supervisor / routinely requires instruction / routinely requires prompting / requires prompting significant correction of work</td>
</tr>
<tr>
<td>Knowledge and understanding of the law / legal practice issues; rarely draws on appropriate prior knowledge or legal principles</td>
<td>Poor knowledge and understanding of law / legal practice issues; rarely draws on appropriate prior knowledge or legal principles</td>
</tr>
<tr>
<td>Strength of oral communication skills</td>
<td>Poor oral communication skills indicating enduring difficulties in articulating legal and factual material; regularly fails to plan, listen or adapt to the needs of the audience</td>
</tr>
<tr>
<td>Strength of written communication skills</td>
<td>Poor written communication skills; rarely shows clarity, precision and accessibility; drafts routinely require significant amendment</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Strength of research skills</td>
<td>Poor research skills; rarely shows appropriate depth, detail and comprehensiveness; reports rarely display effective practical awareness and application</td>
</tr>
<tr>
<td>Commitment to clients and the Student Law Office</td>
<td>Demonstrates little commitment or enthusiasm for achieving the best solution for clients; rarely puts more than the minimum required to perform tasks; completes insufficient work</td>
</tr>
<tr>
<td>Case management and strategizing</td>
<td>Cases are progressed poorly; very few ideas about cases are offered or are poorly formed and not thought through; there is little or no evidence of proactivity or thinking about the overall strategic direction of clients' cases</td>
</tr>
<tr>
<td>Organisation: time and file management</td>
<td>Displays poor organisational skills; makes little effective attempt to manage time; regularly fails to anticipate how long tasks will take or to plan use of time effectively; late on more than three occasions; files are often disorganised and not up to date; copes poorly under pressure and fails to achieve results when time is of the essence</td>
</tr>
<tr>
<td>Teamwork skills and contribution to firm meetings</td>
<td>Poor working relationship with Supervisor / partner / peers; ineffective or negligible or disruptive contribution to firm meetings; may sometimes fail to attend firm or other meetings; relies heavily on other people to achieve client goals</td>
</tr>
<tr>
<td>Understanding of client care and professional conduct</td>
<td>Displays a poor understanding of professional obligations; fails to take client care procedures seriously or fails to ascertain the appropriate office procedure; commits a significant breach of the Code of Conduct or error of professional judgment</td>
</tr>
</tbody>
</table>
## Appendix B
### Reflections matrix Student Law Office

<table>
<thead>
<tr>
<th></th>
<th>Third/fail (below 50)</th>
<th>Lower Second (50-59)</th>
<th>Upper Second (60-69)</th>
<th>First/strong first (70+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflective Analysis</td>
<td>No significant analysis or reflection on the topic</td>
<td>Fair analysis and reflection on the topic. Using some detailed examples but primarily descriptive with a lack of development or analysis.</td>
<td>Good analysis and reflection on the topic. Specific and personal, using some detailed examples, showing good ability to synthesise and evaluate information and ideas</td>
<td>Excellent relevant in depth analysis and reflection on the topic. Specific and (where appropriate) personal, using detailed examples showing excellent ability to synthesise and evaluate information and ideas</td>
</tr>
<tr>
<td>(Self) Awareness and insight (where appropriate, dependent on the topic) *</td>
<td>Exhibits little or no self-awareness, generalises experiences, fails to take into account other perspectives or examine potential value</td>
<td>Exhibits fair/reasonable levels of self-awareness, but some generalisation of experiences, sometimes takes into account other perspectives and examines potential value</td>
<td>Exhibits good levels of self-awareness, avoids generalisation of experiences, often takes into account other perspectives and examines potential value</td>
<td>Exhibits high/very high levels of self-awareness, avoids generalisation of experiences, always takes into account other perspectives and examines potential value. Evidence of development/learning and future development/learning needs</td>
</tr>
<tr>
<td>Context (Knowledge of relevant material and sources)</td>
<td>No evidence of relevant knowledge or independent reading.</td>
<td>Little evidence of relevant knowledge. Relies solely on personal anecdote.</td>
<td>Some evidence of independent reading such as books or journal articles.</td>
<td>Good/Excellent evidence of independent reading such as books or recent journal articles which supports the reflection and or provides context</td>
</tr>
<tr>
<td>Clarity of expression</td>
<td>Not always clear what was intended. Very poor style. Extensive grammar or vocabulary errors</td>
<td>Some points may not be expressed clearly. Poor style. A number of grammar or vocabulary errors.</td>
<td>Most points expressed clearly and succinctly. Mainly engaging and comprehensible style. Mainly correct grammar and vocabulary</td>
<td>All points expressed clearly and succinctly. Engaging and comprehensible style. Correct grammar and vocabulary</td>
</tr>
<tr>
<td>Organisation</td>
<td>Little or no organisation of the material</td>
<td>Clear organisation of material but at times the transitions are unclear.</td>
<td>Very clear organisation of material.</td>
<td>Excellent organisation of the material, forming a coherent whole.</td>
</tr>
</tbody>
</table>

*This may be slightly less relevant in some of the optional titles, such as Clinic and Legal Education*