



## Marking consistency

Ensuring agreement between a team of markers (inter-rater reliability) and that marking is consistent over time (intra-rater reliability) is difficult. A key requirement is the development of a shared understanding of the mark scheme. Whilst this may seem obvious, pre-marking preparation is an activity that often does not receive the time and attention necessary for marker agreement. It may be assumed that everyone 'knows' what work of a particular level or quality looks like, but research has demonstrated that this is not the case. In addition, not only do the mental representations of work 'typical' of particular grades differ, but markers may include features and qualities that are not in fact specified in the mark scheme. Using heuristics, cognitive short-cuts such as 'recognising' a response, may therefore contribute to a lack of marking consistency.

Ensuring a shared understanding of the mark scheme is also important as research has suggested that markers demonstrate confirmation bias, making an initial judgement early in their engagement with a response, which they then seek to confirm. It is therefore important to ensure that this initial judgement is accurate and mark scheme informed and not based on personal and idiosyncratic criteria.



Marking is easy

Experienced markers are better markers

Markers have a shared understanding of the mark scheme

## Steps to enable marking consistency within a marking team

### Understand the marking criteria

- Develop mark schemes which apply discipline-specific meanings to the generic terms used in the University grade descriptors
- All staff (and students) need to familiarise themselves with both versions.

### Convene a pre-marking meeting

- Put egos to one side...Be prepared to change your personalised and possibly idiosyncratic criteria.
- Discuss the marking criteria. A shared understanding may be achieved using a consensual approach, in which agreement is reached through discussion, or a hierarchical approach, in which the module convenor conveys his/her expectations and understanding of different levels of achievement.

### Standardise marking

- One suggestion is for all markers to mark a response or a range of responses prior to an initial marking meeting. Alternatively, responses might be marked during the pre-marking meeting.
- The marked responses provide the basis of discussion for where and why marks were allocated.

### Use exemplars

- Exemplar responses help to 'translate' the marking criteria into marking practice. Best practice is to have a range of exemplars available.
- Exemplars may be used throughout the marking process. Initially to assist with developing a shared understanding of the marking criteria and later as anchor points to refer back to as the marking period progresses. They may also be referred to if discrepancies arise between first and second markers.

### Apply an 'internalised marking schema'

- Embed an accurate understanding of the marking criteria early in the marking period.
- Having a mark scheme informed 'internalised marking schema' will help to protect against the application of idiosyncratic criteria and subjective judgement.

### Feedback

- **Feedback to markers:** Early in the marking period, markers might send examples of marked assignments to the module convenor for verification.
- **Feedback provided on the response:** Use the language of the marking criteria to provide feedback. This will help to ensure that marks are being assigned based on the marking criteria and not for non-content related material or skills that are not specified on the mark scheme.

#### Further reading

- Bloxham, S., Boyd, P., & Orr, S. (2011). Mark my words: the role of assessment criteria in UK higher education grading practices. *Studies in Higher Education, 36*(6), 655-670.
- Code of Practice for Assessment and Feedback (University of Surrey): <https://surrey-content.surrey.ac.uk/sites/default/files/2019-11/Code%20of%20practice%20for%20assessment%20and%20feedback%202019-20%20final.pdf>
- 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.
- Hack, S. (2020). *How do examiners mark? An investigation of marking processes used in the assessment of extended written responses*. (Doctoral dissertation, University of Surrey, Guildford, UK.) Retrieved from <http://epubs.surrey.ac.uk/853450/1/PhD%20thesis%20Sarah%20Hack.pdf>
- Hay, P., & Macdonald, D. (2008) (Mis)appropriations of criteria and standards-referenced assessment in a performance-based subject. *Assessment in Education: Principles, Policy & Practice, 15*(2), 153-168.
- Sadler, D.R. (2013) Assuring academic achievement standards: from moderation to calibration. *Assessment in Education: Principles, Policy & Practice, 20*(1), 5-19.