

## Trialling two proposed grading models for the International Project Qualification

Ben Smith, Robin Drummond & Martin Taylor

### Abstract

This paper describes a trial of two proposed grading models for the International Project Qualification, replacing a more typical marking model. The purpose of this trial was to compare the performance of a novel model with 2 grade descriptors (2-grade model) with a more traditional model with 4 grade descriptors (4-grade model), in terms of grading accuracy and ease of use. The mechanics of the trial are detailed, before its findings are discussed. In general, evidence suggests that the 2-grade model performs marginally better (or at the least, no worse) than the 4-grade model.

### Introduction

During the development of the International Project Qualification (IPQ; intended to be equivalent to the Extended Project Qualification or EPQ in England), CERP researchers identified a number of issues with the EPQ approach, which was used as a starting point for development of the IPQ. One major issue was the EPQ's model of allocating of marks to the project, which resulted in regression of marks during moderation that the development team deemed unfair, as a candidate could drop below a boundary despite the grading of their project having been accurate. As an alternative, it was decided that a model in which projects were graded rather than marked would be investigated.

However, whilst developing such a model, it became clear that writing descriptors which identified distinct levels of performance for *every* element of the project at 6 grades (A\*-E) was not feasible. Nor, indeed, was it easy to generate descriptors for 4 grades (A\*, A, C and E), which was attempted in lieu of describing every grade. After much deliberation, a novel 2-grade model was settled upon as a potential solution to the above problem.

### The 2-grade model

Initially, a grid was drawn up separating out what each assessment objective (AO) element would be likely to look like at each grade (see Appendix 1 for a list of AOs and their weightings, and the first column of tables in Appendix 2 for a breakdown of AO elements). Importantly, the description of AO element performance need not differ between all grades across every AO - it might be identical between A\* and A, be different for a B, then be different again for grades D-E. As an example, see the grid for one AO in Table 1.

**Table 1:** Example grid (AO3: Demonstration of Research Skills)

	A*	A	B	C	D	E
Justify the selection of sources used (3.1)	<b>Sources are selected for a clear and justified purpose.</b>		Most sources are selected for a clear and justified purpose.	<b>Sources are selected, though justification of source choice is sporadic and not always clear.</b>		Source selection is rarely justified
Identify relevant information from sources (3.2)	<b>Relevant information from sources is identified systematically and effectively.</b>		Selection of information from sources is clear, but not systematic and thus may not always be effective.	<b>Information is identified but organisation of it may be deficient and the relevance may not always be clear.</b>		Selection of information from sources is basic
Subject sources to scrutiny (3.3)	Sources are subject to rigorous scrutiny to assess their validity.	<b>Sources are subject to scrutiny to assess their validity.</b>		Scrutiny of sources is generally present and carried out effectively, at least in some instances.	<b>Source scrutiny is brief and not always present.</b>	

The construction of grids as seen in Table 1 was crucial, as it allows one to identify which elements of the AOs performance would be expected to differ from one grade to the next, and which it would not be expected for. See Appendix 2 for the complete grids.

From the grids for each AO, two grade descriptors were constructed, for grades A and D (see Appendix 3). These descriptors included the emboldened red text for each element of each AO, as seen in Table 1 (as an example). The intent of this model is that these two descriptors will be sufficient to grade every project via a 2-stage process. The first stage is, paraphrasing somewhat: “Is the project more like the A or the D descriptor?”, and the second: “Is the project the same as the grade descriptor, or better/worse enough to warrant a change of grade?” (thereby allowing the allocation of the non-described grades). To facilitate the second stage, several ‘typical distinguishing feature’ statements were produced. For example, a marker might decide: a) that a project is closer to the A descriptor than the D descriptor, and b) that the project is sufficiently weaker than the A descriptor to warrant the award of a B grade. A strength of this approach was that, as long as the (relatively simple) initial A/D decision was correct, it would be very difficult for an assessor to ever grade a project incorrectly by more than a single grade (in theory at least). See Appendix 3 for the grade descriptors, accompanying instructions, and ‘distinguishing features’ guidance statements.

However, the grid and the two grade descriptors (with accompanying guidance) do not suitably define how projects pass or fail (and thus how they can receive a grade U). To this end, a series of definitive yes/no statements was developed, which listed all the pieces of work and tasks that were necessary to have a complete project and be eligible to attain any pass grade. The intent was that, if any of the list of statements was answered “no”, then the project would receive a grade U. See Appendix 4 for this list of yes/no statements.

### The 4-grade model

However, the developers had doubts about the viability of the 2-grade model. Despite many of these concerns being debated and resolved, it was decided that it would be wise to trial both the 2-grade model and one with a more traditional set of grade descriptors. This second model had

4 grade descriptors as described earlier (A\*, A, C and E), and followed a more traditional approach to grading based on “which descriptor does the project best fit” (with some additional detail on how to award B and D). These descriptors were akin to those in the existing EPQ specification, where three grade descriptors for grades A, C and E exist, except that the development team had worked to remove as much subjectivity as possible from the descriptors to try to improve the consistency of grading. It is important to note, however, that the grade descriptors in the EPQ are not used for assessment (assigning projects grades), and are merely included for users’ information.

The one other change to this model from the EPQ grade descriptors was the introduction of ‘gating’ statements. Gating requires that certain elements of lower grade descriptors be met, otherwise the candidate cannot access a given grade. The intent of introducing these statements was to, again, make the assessment criteria less subjective than in the EPQ. The 4-grade model’s grade descriptors can be found in Appendix 5.

## **Trialling the models**

### **Design**

The lead IPQ developer provided the names of 10 current EPQ moderators who were contacted and who agreed to take part in the trial. She also examined 19 EPQ projects that demonstrated a wide range of performances, and re-graded them according to her ‘holistic’ sense of the grade they deserved in the IPQ, bearing in mind the differences between EPQ and IPQ AOs and requirements. The ten moderators were split randomly into two groups of five, with each group allocated either the 2-grade or 4-grade model to grade the 19 projects. Both sets of moderators received an identical copy of the AOs and identical instructions on the task they were to undertake. Given the nature of the 2-grade model, how a certain grade was to be arrived at based upon the sequence of decisions (i.e. a) does it pass; b) it is more like an A or a D; c) is it above or below that grade) was spelt out, as is evident in Appendix 3.

### **Quantitative findings**

Table 2 shows the results of this trial in terms of grade differences (away from the lead developer grade). Positive values indicate that the moderator allocated a grade above that of the lead developer, negative below. Note that differences of a fraction of a grade indicate that the moderator could not decide between two grades (they allocated a B/C, as an example). Cells are shaded according how this grade would be rated during live moderation: green for exact agreement; amber for a single grade’s deviation; red for two or more grades’ deviation. For convenience, the sum of absolute deviations on each project and for each marker are also presented in the bottom row and rightmost column.

**Table 2:** Grade differences summary

Project number	Lead developer grade	Moderator number										Sum of absolute deviations on project
		2-grade model					4-grade model					
		1	2	3	4	5	6	7	8	9	10	
1	A*	-1	0	-1	0	-3	0	0	-1	0	-3	9
2	A*	0	0	0	0	0	0	0	0	0	0	0
3	A*	0	-1	-1	0	0	0	-1	-1	-1	-3	8
4	A	0	1	0	1	1	0	1	1	1	-1	7
5	A	-1.5	-2	-1	0	-1	-1	-1	0	-1	0	8.5
6	A	0	0	1	0	0	2	0	-1	0	-2	6
7	B	-0.5	1	1	0	1	0	0	1	1	-1	6.5
8	B	-1	0	-1	-2	-1	0	-1	-2	-1	-2	11
9	B	0	0	0	-1	0	1	-1	-1	0	-1	5
10	C	-1	-1	0	-1	-1	0	-1	-1	-1	-2	9
11	C	1	0	1	1	1	2	-1	-1	1	-1	10
12	C	0	-2	0	0	0	0	-2	-1	-1	-1	7
13	D	0	0	1	0	2	0	-1	0	-1	0	5
14	D	0	1	0	1	0	2	0	2	0	-2	8
15	E	1	3	2	1	0	3	0	2	2	1	15
16	E	0	0	1	1	0	1	2	1	0	0	6
17	E	1	1	1	1	1	-1	0	-1	1	0	8
18	U	0	0	1	1	0	0	0	0	0	1	3
19	U	1	0	1	1	1	1	0	1	0	0	6
<b>Number of exact agreements for moderator</b>		10	10	6	8	9	10	9	4	8	6	
<b>Sum of absolute deviations for moderator</b>		9	13	14	12	13	14	12	18	12	21	

Numerous conclusions can be drawn from this table. First, that there are both more total deviations from the lead developer grade for the 4-grade model (52/95 graded allocated deviate by at least one grade in the 2-grade model; 58/95 do in the 4-grade model) and these deviations tend to be larger in size (average absolute deviation for the 2-grade model was 1.17 grades; for the 4-grade model, 1.33 grades). Note that even if the most erroneous moderator, moderator 10, is removed, there is still a higher proportion of deviations and a higher average absolute deviation for the 4-grade model (1.24 grades) than for the 2-grade model, though the gap between the two models narrows.

It is also interesting to note that certain projects were consistently graded (regardless of model) either higher or lower than the lead developer grade. Project 5, 8, 10, 11 and 15 are the most obvious examples of this, and it is impossible to say whether this is down to an inaccurate lead moderator grade or a consistent error on the part of the moderators. Further, some projects seem better marked under one model or another: projects 3, 6, 9, 12 and 14 have fewer deviations under the 2-grade model, whilst projects 18 and 19 have fewer under the 4-grade model. This raises an important point - the 2-grade model seems, overall, to perform slightly

better in the majority of the mark range, but at the bottom (grades E and U the 4-grade model performs better).

As alluded to with the mention of moderator 10's inaccuracy earlier, there are also trends in terms of moderator accuracy. The moderators for the 2-grade model appear, overall, to be grading more accurately than those using the 4-grade model. The average number of exact agreements (per moderator) in the 2-grade model is on average, one higher than in the 4-grade model, 8.6 as opposed to 7.4. Looking at 'Sums of absolute deviations for moderator', the 2-grade model gain appears to be more reliable, largely down to the high overall deviation rate of moderators 8 and 10. Unfortunately the small number of moderators used means it is not possible to determine whether the 4-grade model's seeming poorer performance was down to less accurate moderators being randomly selected to trial it. However, it is at least reassuring that the best performing moderators in the 4-grade model appear similar to those in the 2-grade model. From the available data, there is thus little evidence that the 2-grade model performs worse than the 4-grade one (though admittedly the validity of the trial is called into question by the small and potentially incomparable groups of moderators using each model).

In addition, it is worth noting that a slightly under 50% exact agreement rate across all grading is similar to the proportion of projects in the EPQ which are allocated exactly the same grade by the moderator and the centre. Whilst the individuals grading here are experienced moderators, they were doing so on a completely new model with very little instruction and no standardisation. This could, therefore, be taken to indicate that a grading model may result in similar degrees of grading inaccuracy to a typical marking one - though admittedly drawing this conclusion requires some leaps in logic.

### **Qualitative findings**

As is evident, the quantitative data in Table 2 is not from a large enough sample to draw reliable conclusions about which model performed better. However, the ten moderators taking part in the trial were each asked to provide qualitative "feedback about how easy you found the grading process to follow and any suggested improvements" via an email questionnaire, to be recorded during and immediately after the grading process. It is important to bear in mind that these moderators were accustomed to the EPQ process of marking by AO and summing these marks to produce an overall mark - not grading in a holistic manner. One moderator for the 2-grade model (moderator 3) did not provide comments of sufficient length to draw any conclusions, thus his/her feedback is not considered here. The qualitative analysis conducted was brief and did not fully conform to any formal technique – thematic analysis is closest, however, as the similar ideas and concepts emerging from multiple moderators' commentaries comprise the data pulled out and discussed below.

In general, the moderators of the 2-grade model acknowledged that initially, using this model was challenging as it was so much of a departure from what they had done before with the EPQ. Two of the moderators had, however, decided by the end of the process that they liked the model and could work with it in future. The other two stated their reservations as wanting more detail on fringe cases (which the grid (Appendix 2), not provided to them in the trial, provides in greater detail than the additional guidance in Appendix 3) and having concerns about 'best fit' descriptors (which could apply to either model).

The moderators of the 4-grade model were again fairly split as to their preference for this model or more traditional assessment criteria (using marks). Two of the five moderators liked this model and would have preferred it to a mark-based one, the others did not. Note that the reasons for preferring it were largely that the approach is holistic, which applies to both the 2-grade and the 4-grade models. However, four of the five moderators (barring one of the two moderators who liked the 4-grade model) all noted a fairly large number of issues that tended to be common across at least two moderators. A lack of clear links to the AOs caused difficulty in

weighting different parts of the assessment during holistic judgements, and the intermediate grades without descriptors (D and B) proved especially troublesome. Further, a number of moderators identified an issue with how tightly defined some of the statements in the grade descriptors were, together with 'gating' statements: that candidates must meet all the below grade criteria to access higher grades. The moderators expressed concern about whether or not they were supposed to apply compensation given that 'gating' statements were present. Moderator 10 seems to have taken this to extremes, evidenced by his/her consistently severe grading (Table 2), applying the gating rigidly rather than using a more holistic compensatory approach that other moderators employed (which, on the basis of the lead developer grades, would seem to be the intended approach).

Both sets of moderators acknowledged that due to their pre-existing knowledge of the EPQ, they were unlikely to be able to use the IPQ grading models independently. It is unknown whether this may have had a negative, positive, random, or negligible impact on the quality of their grading in this study.

## Discussion

The sample is too small to draw reliable conclusions from the quantitative evidence. That said, in general, the 2-grade model does appear to perform marginally better, from the quantitative evidence. This is both in terms of the proportion of exact agreements and the magnitude of disagreements with the lead examiner's grade (and the better performance of the 2-grade model persists when the most erroneous 4-grade model moderator is removed). However, it is of course possible that this outcome is entirely down to the better moderators being selected into the 2-grade model by random chance. Whilst it is therefore not possible to conclude, based upon this small sample, that the 2-grade model performs better than the 4-grade, it is reasonable to draw some assurance that, at the very least, there is no evidence that the novel approach of the 2-grade model is *worse* than a more traditional set of grade descriptors (the 4-grade model). There is also some evidence to suggest that either model results in a similar proportion of moderator-lead developer grade agreements as there are centre-moderator grade agreements in the IPQ at present.

The one aspect of the quantitative evidence that is concerning is that at grades E and U, the 4-grade model seems to perform better than the 2-grade model. In order to address this shortcoming, the 2-grade model was amended to feature additional E/U 'typical distinguishing features' statements, to be utilised if a project is identified as being below D to double-check whether it is worthy of an E or not (see Appendix 6). Note that this shifts the initial Yes/No criteria from being necessary and sufficient criteria for a pass, to merely necessary criteria for a pass.

The qualitative evidence is interesting, but harder to delineate. All moderators acknowledged that they may be biased due to prior exposure to the EPQ, and both sets of moderators had mixed opinions of their model, some positive and some negative. However, the difference is in what aspects of the models were praised and critiqued by the respective moderators. The two grade model's innovative approach was liked by several, and crucially, the issues that were raised could either: be addressed through use of the grid (giving more detail when judging projects falling between two grades); or applied equally to both models (concerns about best fit descriptors). Similarly, the aspects of the 4-grade model considered to be positive could equally be applied to the 2-grade model, as they were features of *any* holistic grading model. For the 4-grade model a number of issues were raised that were specific to the design of that model (particularly a lack of clear links to AOs and gating statements). Whilst the moderator feedback for each model was mixed, it is reasonably clear that there were few identified positives of the 4-grade model and few identified negatives of the 2-grade model that were not common to both models. This does seem to point to the 2-grade model being a preferable option - though of

course, no moderators had both models available to them, so this conclusion is perhaps not entirely fair. Again though, whilst one cannot say that the qualitative evidence points to the 2-grade model definitely being superior to the 4-grade one, it is certainly reasonable to say that the 2-grade model does not seem any more problematic than the 4-grade one.

It is also worth considering one final point, namely that there were numerous differences between the 2-grade and the 4-grade model. These included the method of awarding grades (a stepped process in the 2-grade model; a less defined best-fit process in the 4-grade model) and the clarity of grade descriptors (as discussed above, gating statements in particular were identified as issues with the 4-grade model's descriptors). As such, even if it were to be concluded that these findings mean one model is better than the other, it is not possible to delineate precisely the elements of the superior model that resulted in an improvement to grading accuracy. This is important to bear in mind, though it is admittedly inevitable in a test that does not make use of control conditions.

In practice, these conclusions were taken as sufficient evidence that the 2-grade model, if not clearly the better option, would at least seem likely to function no worse than the 4-grade model (and also potentially to result in no greater grading inaccuracy than a marking system). As such, the IPQ is set to utilise the 2-grade model (with minor amendments, notably including the addition of E/U information) as the means to determine grades for each project. Since this model has been adjusted following the trial, the revised yes/no statements, grade descriptors and 'distinguishing features' statements are included in Appendix 6 for reference. With the first submission of projects for the IPQ scheduled in early 2019, it will be some time until the assessment's functioning in a live environment can be established, though there may be further trials (likely comparing a grading to a marking process) in the interim.

## Appendix 1 - IPQ Assessment Objectives

### Assessment Objectives (AOs)

All candidates are required to meet the following Assessment Objectives. The Assessment Objectives are weighted as indicated below:

	<b>Assessment Objectives</b>	<b>Weighting</b>
<b>Selection of Topic AO1</b>	Identify and select an area of interest, devise aims and objectives and establish a working title using a selection of appropriate sources.	10%
<b>Planning, Monitoring and Developing AO2</b>	Produce a project plan which incorporates a time schedule, undertake a risk assessment and monitor progress against objectives.	20%
<b>Demonstration of Research Skills AO3</b>	Research, select, organise and scrutinise a range of information and resources, justifying the selection of any secondary and primary sources used.	20%
<b>Analysis and Application of Research AO4</b>	Implement the project plan, analyse the findings of the research and synthesise these findings into a 5000 word report with references and bibliography. Include a conclusion and communicate outcomes in both a report and a presentation.	40%
<b>Evaluation of Product, Process and Self AO5</b>	Evaluate the strengths and weaknesses of the project process (including review of own learning arising from it) and product.	10%



## Appendix 2 - 2-grade model AO grids

### A01 – Selection of Topic

	A*	A	B	C	D	E
Identify and select an area of interest (1.1)	<b>An area of interest is identified and selected with careful consideration</b>			<b>An area of interest is selected with some care</b>		
Set aims and objectives (1.2)	<b>Aims and objectives are set clearly and precisely</b>			Aims and objectives are set clearly, though precision may be lacking (11)	<b>Precision and clarity may be lacking in the aims and objectives.</b>	Aims and objectives are communicated in broad statements (16)
Develop a working title by researching appropriate sources (1.3)	The working title is well-developed and informed by highly effective preliminary research (1)	<b>The working title is well-developed and informed by effective preliminary research</b>	Development of the working title is supported by some preliminary research (a) and may not always be effective (6)	<b>The working title is developed to some extent, though support from preliminary research may be limited</b>		

**A02 - Planning, Monitoring & Developing**

	<b>A*</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Produce a project plan with a time schedule (2.1)	Planning is carried out with a meticulous level of detail (2)	<b>Planning is detailed and a clear strategy is evident throughout. (b)</b>		<b>Some planning is present but it may lack strategy and detail.</b>		There is an outline plan (17)
Undertake a risk assessment (2.2)	<b>A thorough and well thought out risk assessment is carried out, with awareness of risk evident throughout the project</b>		A risk assessment is present and is detailed. In addition, there may be some awareness of risk elsewhere in the project.	<b>A risk assessment is present but may be superficial, and there is unlikely to be awareness of risk elsewhere in the project</b>		
Select, justify and use monitoring techniques to monitor progress against the plan (potentially using novel project management techniques) (2.3)	<b>Monitoring is comprehensive, and justification of the monitoring method selected is sound.</b>		Monitoring is detailed (although there may be gaps), and justification of the monitoring method is present (but not necessarily sound) (12) (7)	<b>Monitoring is present but there may be gaps, and justification of the monitoring method may be lacking.</b>		
Develop the project by making changes to the plan and working title, justifying decisions made (2.4)	Project development is innovative and rigorous (2)	<b>Project development is detailed, clear and justified throughout. (b)</b>		<b>Project development is clear but may lack detail and justification in places.</b>		

**A03 – Demonstration of Research Skills**

	<b>A*</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Justify the selection of sources used (3.1)	<b>Sources are selected for a clear and justified purpose.</b>		Most sources are selected for a clear and justified purpose. (c)	<b>Sources are selected, though justification of source choice is sporadic and not always clear.</b>		Source selection is rarely justified (18)
Identify relevant information from sources (3.2)	<b>Relevant information from sources is identified systematically and effectively.</b>		Selection of information from sources is clear, but not systematic and thus may not always be effective. (8)	<b>Information is identified but organisation of it may be deficient and the relevance may not always be clear.</b>		Selection of information from sources is basic (18)
Subject sources to scrutiny (3.3)	Sources are subject to rigorous scrutiny to assess their validity. (3)	<b>Sources are subject to scrutiny to assess their validity.</b>		Scrutiny of sources is generally present and carried out effectively, at least in some instances (13)	<b>Source scrutiny is brief and not always present.</b>	

**AO4 - Analysis and Application of Research**

	<b>A*</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Implement the project plan (4.1)	The project plan is implemented clearly, rigorously and with a high level of sophistication.	<b>The project plan is implemented clearly throughout.</b>		<b>A clear attempt has been made to implement the project plan, though this may not have always been effective.</b>		
Produce a 5000 word report addressing the selected final title in an appropriate writing style (with focus on title) (4.2)	<b>Report is of appropriate length and writing style, with a clear focus on the final title throughout.</b>		The report is of an appropriate length and writing style, but its focus is not maintained throughout (14)	<b>The report may lack length, focus or an appropriate style.</b>		The report is somewhat disorganised and lacks focus.
Organisation and referencing in the report and the presentation (4.3)	The report and presentation are organised well and referencing is complete and with minimal error.	<b>The report and presentation are organised well and referenced appropriately.</b>		<b>The report and presentation show some organisation, but information and/or sources may not be referenced appropriately.</b>		The report and presentation are somewhat disorganised and lack focus (19)
Analysis of the research findings in the report and the presentation (4.4)	Analysis has a degree of nuance that demonstrates subject mastery. (4)	<b>Analysis is thorough</b>		<b>Analysis may lack depth and there may be gaps in the analysis</b>		
Synthesise the selected information and sources into the report and draw a conclusion (4.5)	<b>Selected information and/or sources are synthesised into a compelling report that leads to a well-evidenced conclusion.</b>		The report is compelling, (d) and a conclusion is present but evidence may be lacking (9)	<b>There is an attempt to synthesise selected information and/or sources into the report but the resulting conclusion may not be well-evidenced.</b>		
Communicate findings in the report and the presentation (4.6)	<b>Findings are communicated clearly and reflect a sound grasp of the subject matter.</b>		Findings are communicated clearly (d)	<b>Communication of findings may be slightly unclear at times.</b>		

**A05 - Evaluation of Product, Process & Self**

	<b>A*</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Assess the strengths and weaknesses of the final report, including in relation to aims and objectives (in the Production Log and presentation) (5.1)	Evaluation is detailed, insightful and sophisticated (5)	<b>Evaluation is detailed and insightful</b>	Evaluation is detailed (e) but lacks insight (10)	<b>Some evaluation but it may lack detail and/or not fully consider all aspects of the project.</b>		Evaluation is simplistic (20)
Assess the strengths and weaknesses of the organisation of the project (in the Production Log and presentation) (5.2)	Evaluation is detailed, insightful and sophisticated (5)	<b>Evaluation is detailed and insightful</b>	Evaluation is detailed (e) but lacks insight (10)			Evaluation is simplistic (20)
Review own learning from the process of completing the project (in the Production Log and presentation) (5.3)	Evaluation is detailed, insightful and sophisticated (5)	<b>Evaluation is detailed and insightful</b>	Evaluation is detailed but lacks insight (10) (15)	<b>Evaluation may lack detail and/or not fully consider all aspects of the student's own learning.</b>	Evaluation is simplistic (20)	

### Appendix 3 - 2-grade model grade descriptors & 'distinguishing features' statements

#### Which descriptor is this project most similar to?

This should be done as a best fit looking at the characteristics of the grade D and then the grade A to see which is most appropriate. If having difficulty, you can examine the B/C borderline guidance below to see if this helps decide whether it is more similar to a D or more similar to an A.

A	D
<b>Evidence towards the following descriptions will be found throughout the project, particularly in the three necessary elements: the Production Log, the Report and the Presentation.</b>	
<p>The student <u>carefully considers and selects</u> an appropriate area of interest and sets aims and objectives <u>clearly and precisely</u>. The working title has been <u>well-developed</u> and is informed by effective preliminary research.</p> <p><u>Planning is detailed</u> and a <u>clear strategy</u> is evident throughout. There is <u>comprehensive</u> monitoring of progress against the plan, and the selection of a monitoring technique is <u>clearly justified</u> (monitoring may be done using a novel project management technique). Project development is <u>detailed, clear and justified throughout</u>.</p> <p>Sources are selected for a <u>clear and justified purpose</u> and relevant information from them is identified <u>systematically and effectively</u>. Sources are subjected to <u>scrutiny to assess their validity</u>.</p> <p>The project plan is <u>implemented clearly</u> throughout. The report is of <u>appropriate length and writing style</u>, with a <u>clear focus on the final title throughout</u>. The project is <u>organised well</u> and referenced appropriately. Analysis of the research findings is <u>thorough</u>. Selected information and/or sources are <u>synthesized into a compelling report</u> that leads to a <u>well-evidenced conclusion</u>. Findings are <u>communicated clearly</u> in both the report and the presentation and reflect a sound grasp of the subject matter.</p>	<p>The student identifies and selects an appropriate area of interest with <u>some care</u> and sets aims and objectives, <u>though precision and clarity may be lacking</u>. There is <u>some development</u> of a working title, though support from preliminary research may be limited.</p> <p><u>Some planning</u> is present, but it may lack strategy and detail. There is monitoring of progress against the plan, but <u>there may be gaps</u> and the selection of a monitoring technique may <u>lack justification</u>. Project development is <u>clear but may lack detail and justification in places</u>.</p> <p>Sources are selected, though <u>justification of source choice is sporadic</u> and not always clear. Information from sources is identified <u>but organisation of it may be deficient</u> and the <u>relevance may not always be clear</u>. Source <u>scrutiny is brief</u> and not always present.</p> <p>A <u>clear attempt</u> has been made to implement the project plan, though this may not have always been effective. The report may <u>lack length, focus or an appropriate style</u>. The project shows some organisation, but information and/or sources may not be referenced appropriately. Analysis of the research findings may <u>lack depth</u> and there may be gaps in the analysis. There is an <u>attempt to synthesize</u> selected information and/or sources into a report, but the resulting conclusion <u>may not be well-evidenced</u>. Communication of findings in the report and the presentation <u>may be unclear at times</u>.</p>

<p>There is <u>detailed and insightful evaluation</u> of the strengths and weaknesses of both the final report and the organisation of the project. The student reflects on his/her own learning in a <u>detailed and insightful</u> manner.</p>	<p>There is <u>some evaluation</u> of the strengths and weaknesses of both the final report and the organisation of the project, but this may <u>lack detail</u> and/or not fully consider all aspects of the project. The student reflects on his/her own learning to some extent, though this may lack depth.</p>
--	---

In the event that it is unclear whether a project is most similar to the D or to the A descriptor (and therefore whether it is likely to achieve a C or a B), the following guidance can be used to aid your decision.

Typical features of a B that a C may not have:

- AO1: The development of the title is supported by preliminary research
- AO2: Planning and development are detailed and strategic
- AO3: Most sources are selected for a clear and justified purpose
- AO4: The report and presentation communicate findings clearly and discussion is compelling
- AO5: Evaluation is detailed throughout

**Does the project match the chosen descriptor (D or A), or is it sufficiently better or worse to warrant a different grade?**

If the descriptor is not met fully then it will be the grade below. If having difficulty, examine the guidance below, which gives some examples of typical features that push a project “over the edge” from one grade to another.

Typical distinguishing features of A\* (from an A):

- AO1: The design of the project is based on highly effective preliminary research
- AO2: The project is planned in a meticulous level of detail and development is innovative and rigorous
- AO3: Sources are subjected to rigorous scrutiny to assess their validity
- AO4: The report analyses findings with a degree of nuance that demonstrates subject mastery
- AO5: Evaluation is insightful and sophisticated throughout

Typical distinguishing features of B (from an A):

- AO1: Preliminary research is present but may not always be effective
- AO2: Monitoring of the project is detailed, although there may be gaps (and novel project management techniques, if present, are likely to lack justification as to their selection)
- AO3: Selection of information from sources is clear, but not systematic and thus is less effective
- AO4: A conclusion is present but evidence for it may be lacking
- AO5: Evaluation of the product and process is detailed but lacks insight

Typical distinguishing features of C (from D):

- AO1: Sets aims and objectives clearly though precision may be lacking
- AO2: Monitoring of the project and the risk assessment are detailed (although there may be gaps)
- AO3: Scrutiny of sources is generally present and is carried out effectively, at least in some instances
- AO4: The report is of appropriate length and writing style
- AO5: Evaluation of own learning is detailed

Typical distinguishing features of E (from D):

- AO1: Project aims are communicated only in broad statements
- AO2: The project plan is more of an outline
- AO3: Source selection is rarely justified and selection of information from sources is basic
- AO4: The report is somewhat disorganised and lacks focus
- AO5: Evaluation is simplistic throughout



## Appendix 4 - 2-grade model yes/no statements

### Has the student met the minimum criteria for a pass (Grade E or higher)?

If there are any "No" answers, the student can only receive a Grade U. You should also make sure that the work in each area is of sufficient quality to justify reaching a basic pass for a Level 3 qualification.

Student identifies and selects an area of interest appropriate for a Level 3 project.	Yes	No
Student develops a working title which will allow them to produce work of a Level 3 standard.	Yes	No
Student produces a project plan.	Yes	No
Student implements the project plan and documents any changes to it.	Yes	No
Student produces a report addressing the selected final title which is at a Level 3 standard.	Yes	No
Student uses a referencing method.	Yes	No
Student communicates findings in the report and the presentation.	Yes	No
Student creates a bibliography/reference list.	Yes	No
Student evaluates the strengths and weaknesses of the project.	Yes	No

## Appendix 5 - 4-grade model grade descriptors & accompanying guidance

Grade	Grade descriptor
E	<p>Spelling, punctuation and grammar are sufficiently accurate so that meaning is clear</p> <p>Style of writing is appropriate to purpose and to complex subject matter</p> <p>Information is clearly and coherently organised, using specialist vocabulary when appropriate.</p> <p>A broad project aim is stated.</p> <p>A title is selected and an outline project plan is produced</p> <p>Progress against the plan is recorded in the Production Log</p> <p>Information is gathered from sources and a bibliography drawn up</p> <p>A referenced report, at level 3 standard, incorporating the researched material is produced which addresses the title and reaches a conclusion</p> <p>The candidate delivers a presentation suitable for a non-specialist audience and can answer straightforward questions on their subject matter and the implementation of their plan</p> <p>Strengths and weaknesses of the outcome and the process are reviewed</p>

Grade	Grade descriptor
C	<p><b>All elements of E descriptor +</b></p> <p>The aim and objectives of the project are stated</p> <p>The scope of the title is developed to fit the 5000-word target</p> <p>The project plan is time-referenced and includes research options and report planning</p> <p>Academic and professional sources are accessed and relevant material selected with reasons given. The possibility of bias/unreliability is considered.</p> <p>A 5000-word referenced report, produced in grammatically accurate and correctly spelled English which evaluates the research findings and reaches a conclusion based on evidence (allow report to be slightly over or under 5000 words)</p> <p>The candidate demonstrates a sound grasp of the project subject matter during their presentation and during the subsequent question and answer session</p> <p>The extent to which the plan was implemented is reviewed.</p> <p>The candidate reflects in depth on their own learning</p>

Grade	Grade descriptor
<b>A</b>	<p><b>All elements of E and C descriptors +</b></p> <ul style="list-style-type: none"> <li>• The aim is clearly stated with precise objectives identified</li> <li>• The title is developed/refined to require evaluation.</li> <li>• The project plan is detailed with clear explanation of any changes that are made as the project develops</li> <li>• There is clear evidence of independent organisation of the implementation of the plan.</li> <li>• A wide range of appropriate sources is accessed and critically analysed, including consideration of bias/reliability</li> <li>• Criteria for selection of sources are set out clearly and applied.</li> <li>• There is detailed analysis of data and interpretation of text.</li> <li>• The title is fully addressed with arguments supported by analysed and validated sources.</li> <li>• The report is logical and fluently written.</li> <li>• The report examines the arguments thoroughly before reaching a conclusion.</li> <li>• The presentation is thoughtfully planned and structured, and its delivery successfully engages the audience</li> <li>• There is a thorough critical review of the extent to which the aims were achieved with an understanding of where improvements might have been made.</li> </ul>

Grade	Grade descriptor
<b>A*</b>	<p><b>All elements of E, C and A descriptors +</b></p> <ul style="list-style-type: none"> <li>• Candidates independently identify an original title which is a consequence of detailed preliminary research and development</li> <li>• Decision-making is of a high order and thoroughly discussed in the Production Log</li> <li>• The work is meticulously planned with attention to detail and this is evidenced at every stage.</li> <li>• There is a thorough understanding shown of the limitations of some sources and effective analysis of possible bias. Data is subjected to rigorous scrutiny before inclusion.</li> <li>• The report is sophisticated and well structured. The candidate demonstrates understanding of the subtleties and complexities of the topic.</li> <li>• The high quality of the writing brings lucidity to explanations and arguments.</li> <li>• The presentation reinforces the subject mastery demonstrated in the report.</li> <li>• The report and the Production Log are evaluative throughout, showing accurate assessment of the strengths and possible weaknesses of both outcome and process.</li> <li>• Aims and objectives are fully achieved.</li> </ul>

Grading is undertaken via these descriptors:

- If the work fails to meet the E descriptor, award grade U.
- If the work meets (largely) the Grade E descriptor, award grade E.
- If the work meets (largely) the descriptor for grade C, award grade C.
- If the work is of higher merit than as described by grade E but does not meet the C descriptor, award grade D.
- If the work meets (largely) the descriptor for grade A, award grade A.
- If the work is of higher merit than that as described by grade C but does not meet the A descriptor, award grade B.
- If the work meets the descriptor for grade A\*, award grade A\*.

## **Appendix 6: Revised 2-grade model**

Before beginning to determine a grade for a project, refer to the following list to ensure all the necessary elements of the project are present. If the student has not produced (and submitted evidence of) any of the listed elements, the project cannot be graded and must be allocated a U.

- Student identifies and selects an area of interest
- Student sets a working title
- Student undertakes a risk assessment
- Student produces a project plan
- Student implements the project plan and documents any changes to it
- Student produces a report addressing the selected final title
- Student uses a referencing method
- Student communicates findings in the report and the presentation
- Student creates a bibliography/reference list
- Student evaluates the strengths and weaknesses of the project

The grading process is outlined in two stages, as follows:

### **Step 1: Which descriptor is this project most similar to?**

This should be done as a best fit looking at the characteristics of the grade D and then the grade A to see which is most appropriate. If having difficulty, you can examine the B/C borderline guidance below to see if this helps decide whether it is more similar to a D or more similar to an A.

Please note that this decision places the project into either the A\*-B range, or C-U range. It is unlikely that any project, once judged to be more like either an A or a D, would slip into the other range during the later steps.

A	D
<p>Evidence towards the following descriptions will be found throughout the project, particularly in the three necessary elements: the Production Log, the Report and the Presentation.</p>	
<p>The student <u>carefully considers and selects</u> an appropriate area of interest and sets aims and objectives <u>clearly and precisely</u>. The working title has been <u>well-developed</u> and is informed by effective preliminary research.</p> <p><u>Planning is detailed</u> and a <u>clear strategy</u> is evident throughout. There is <u>comprehensive</u> monitoring of progress against the plan, and the selection of a monitoring technique is <u>clearly justified</u> (monitoring may be done using a novel project management technique). Project development is <u>detailed, clear and justified throughout</u>.</p> <p>Sources are selected for a <u>clear and justified purpose</u> and relevant information from them is identified <u>systematically and effectively</u>. Sources are subjected to <u>scrutiny to assess their validity</u>.</p> <p>The project plan is <u>implemented clearly</u> throughout. The report is of <u>appropriate length and writing style, with a clear focus on the final title throughout</u>. The project is <u>organised well</u> and referenced appropriately. Analysis of the research findings is <u>thorough</u>. Selected information and/or sources are <u>synthesized into a compelling report</u> that leads to a <u>well-evidenced conclusion</u>. Findings are <u>communicated clearly</u> in both the report and the presentation and reflect a sound grasp of the subject matter.</p>	<p>The student identifies and selects an appropriate area of interest with <u>some care</u> and sets aims and objectives, <u>though precision and clarity may be lacking</u>. There is <u>some development</u> of a working title, though support from preliminary research may be limited.</p> <p><u>Some planning</u> is present, but it may lack strategy and detail. There is monitoring of progress against the plan, but <u>there may be gaps</u> and the selection of a monitoring technique may <u>lack justification</u>. Project development is <u>clear but may lack detail and justification in places</u>.</p> <p>Sources are selected, though <u>justification of source choice is sporadic</u> and not always clear. Information from sources is identified <u>but organisation of it may be deficient</u> and the <u>relevance may not always be clear</u>. Source <u>scrutiny is brief</u> and not always present.</p> <p>A <u>clear attempt</u> has been made to implement the project plan, though this may not have always been effective. The report may <u>lack length, focus or an appropriate style</u>. The project shows some organisation, but information and/or sources may not be referenced appropriately. Analysis of the research findings may <u>lack depth</u> and there may be gaps in the analysis. There is an <u>attempt to synthesize</u> selected information and/or sources into a report, but the resulting conclusion <u>may not be well-evidenced</u>. Communication of findings in the report and the presentation <u>may be unclear at times</u>.</p>

There is detailed and insightful evaluation of the strengths and weaknesses of both the final report and the organisation of the project. The student reflects on his/her own learning in a detailed and insightful manner.

There is some evaluation of the strengths and weaknesses of both the final report and the organisation of the project, but this may lack detail and/or not fully consider all aspects of the project. The student reflects on his/her own learning to some extent, though this may lack depth.

In the event that it is unclear whether a project is most similar to the D or to the A descriptor (and therefore whether it is likely to achieve a C or a B), the following guidance can be used to aid your decision.

Typical features of a B that a C may not have:

- AO6: The development of the title is supported by preliminary research
- AO7: Planning and development are detailed and strategic
- AO8: Most sources are selected for a clear and justified purpose
- AO9: The report and presentation communicate findings clearly and discussion is compelling
- AO10: Evaluation is detailed throughout

**Step 2: Does the project match the chosen descriptor (D or A) so that it is worthy of precisely that grade, or is it sufficiently better or worse to warrant a different grade?**

If having difficulty, examine the guidance below, which gives some *examples* of typical features that may push a project “over the edge” from one grade to another.

If the project was **more like an A**, you may need to utilise the following:

*Typical distinguishing features of A\* (from an A)*

- AO6: The design of the project is based on highly effective preliminary research and facilitates an enterprising investigative project
- AO7: The project is planned in a meticulous level of detail and development is innovative and rigorous
- AO8: Sources are subjected to rigorous scrutiny to assess their validity
- AO9: The report analyses findings with a degree of nuance that demonstrates subject mastery
- AO10: Evaluation is insightful and sophisticated throughout

*Typical distinguishing features of B (from an A)*

- AO6: Preliminary research is present but may not always be effective
- AO7: Monitoring of the project is detailed, although there may be gaps (and novel project management techniques, if present, are likely to lack justification as to their selection)
- AO8: It is made clear why information from sources is selected, but this selection is not systematic and thus is less effective
- AO9: A conclusion is present but evidence for it may be lacking
- AO10: Evaluation of the product and process is detailed but lacks insight

If the project was **more like a D**, you may need to utilise the following:

*Typical distinguishing features of C (from D)*

- AO6: Sets aims and objectives clearly though precision may be lacking
- AO7: Monitoring of the project and the risk assessment are detailed (although there may be gaps)
- AO8: Scrutiny of sources is generally present and is carried out effectively, at least in some instances
- AO9: The report is of appropriate length and writing style
- AO10: Evaluation of own learning is detailed

*Typical distinguishing features of E (from D)*

- AO6: Project aims are communicated only in broad statements
- AO7: The project plan is more of an outline
- AO8: Source selection is rarely justified and selection of information from sources is basic
- AO9: The report is somewhat disorganised and lacks focus
- AO10: Evaluation is simplistic throughout

IMPORTANT: If you judge that the project is worse than a D using the “*Typical distinguishing features of E (from D)*”, you **must** also utilise the guidance below to check whether the project is worthy of a pass (a Grade E) as opposed to a fail (a Grade U).

*Typical distinguishing features of a U (from E)*

- AO1: The working title is stated rather than developed (and may have no support from preliminary research)
- AO2: Planning, monitoring and project development are minimal and/or superficial (and may not directly link to the title/aims/objectives)
- AO3: Source scrutiny either is absent, or is limited and superficial
- AO4: The project plan is poorly implemented;  
Analysis is limited, with a descriptive approach dominant;  
Communication of findings is confused
- AO5: Evaluation either is absent or is superficial/minimal