

IMPROVING PERFORMANCE IN CONTRACTUAL PROBLEM SOLVING: CLARIFYING CRITERIA AND STANDARDS THROUGH THE USE OF EXEMPLARS

JOHN JURIAN SZ* AND DAVID NEWLYN**

ABSTRACT

The implementation of criterion and standards-based assessment regimes in institutions of higher education has been driven by the belief in its capacity to support student learning and in its support of transparent and defensible assessment judgments. However, the provision of criteria and standards alone is insufficiently robust to ensure that students are adequately equipped with the requisite knowledge to satisfy the stated criteria and standards. This paper concerns the first phase of a larger research project which aims to contribute to the development of the effectual use of criteria and standards by the introduction of exemplars in a systematic fashion into the learning resources provided for students engaged in LLB units.

In this phase of the project, students were provided with access to five annotated exemplars of law assignments. These exemplars contained detailed feedback on papers which had received fail, pass, credit, distinction and high distinction grades. Each paper received detailed individualised comments in addition to the various performances being mapped against the articulated criteria and standards. Two questionnaires were developed to test the impact of exemplars on students' experience in the performance of assessment tasks: one during the preparatory phase of assessment completion and one following submission of the assessment. These surveys will be continued over a number of years in an attempt to discern how student learning and teaching practices can be best improved. This paper sets out our preliminary results and outlines our objectives to further refine the use of exemplars.

I. INTRODUCTION

Contemporary legal education has increasingly moved towards the requirement of an express articulation of assessment criteria and standards in an effort to ensure objective legitimacy in the measurement of student performance.¹ This trend towards objective pedagogical transparency has become a hallmark of best practice in many Australian law schools.

Despite these worthy goals, the provision of criteria and standards alone is insufficiently robust to ensure that students are adequately equipped with the requisite skills and knowledge to satisfy the stated criteria and standards. This paper marks the first phase of a larger research project which aims to contribute to the development and implementation of criteria and standards

* John Juriansz is a Lecturer at the University of Western Sydney where he lectures in Contract Law and Equity, Trusts and Remedies. His previous commercial experience includes commercial litigation and mergers and acquisitions work at Mallesons Stephen Jaques and Gilbert + Tobin Lawyers.

** Dr David Newlyn is a Lecturer at the University of Western Sydney where he lectures in Introduction to Business Law and Contract Law.

1 Kathryn Ecclestone, "I Know a 2: I when I See It": Understanding Criteria for Degree Classification in Franchised University Programs' (2001) 25 *Journal of Further and Higher Education* 301; Berry O'Donovan, Margaret Price and Chris Rust, 'Know What I Mean? Enhancing Student Understanding of Assessment Standards and Criteria' (2004) 9(3) *Teaching in Higher Education* 325; Berry O'Donovan, Margaret Price and Chris Rust, 'Developing Student Understanding of Assessment Standards: A Nested Hierarchy of Approaches' (2008) 13(2) *Teaching in Higher Education* 205; D Royce Sadler, 'Indeterminacy in the Use of Preset Criteria for Assessment and Grading' (2009) 34(2) *Assessment & Evaluation in Higher Education* 159.

by the introduction of exemplars in a systematic fashion into the learning resources provided for students enrolled into Contract Law, a compulsory LLB unit at the University of Western Sydney in which the assessments consist of legal problem solving tasks.

Contract Law is a compulsory first-year law unit offered simultaneously across two campuses each autumn semester. For formal assessment purposes, students in Contract Law have traditionally undertaken two items of assessment: a take-home assignment and a formal examination.² This research project examines the formal take-home assignment.

Early in 2009, all students undertaking Contract Law were provided with access to five annotated exemplars of 2008 Contract Law student assignments. These exemplars contained detailed feedback on papers which had received fail, pass, credit, distinction and high distinction grades. In addition to the detailed individualised comments, each of the performances were mapped against the articulated criteria and standards. Two questionnaires were developed to test the impact of exemplars on students' experience in the performance of the major in-term assessment task: one during the preparatory phase of the assessment task and one following submission of the assessment. These surveys will be continued over a number of years in an attempt to discern how student learning and teaching practices can be best improved.

II. CRITERIA AND STANDARDS-BASED ASSESSMENT

During 2006, the School of Law at the University of Western Sydney set out on an enterprise to revamp its assessment regime by developing a criteria and standards-based assessment policy which is to be fully implemented by 2010. This new policy has sought to clarify aspects of the pedagogical and philosophical bases of this newly adopted assessment practice. Fundamental to this new policy have been the loaded concepts of 'criteria' and 'standards'. *Criteria* have been described as the 'specific performance attributes that the assessor takes into account when making a judgment about the student response to the different elements of the assessment task'.³ *Standards* have been defined to refer to 'statements describing the level or quality of student performance in an assessment task'.⁴

III. STUDENT (AND ASSESSOR) COMPREHENSION OF ASSESSMENT CRITERIA AND STANDARDS

Assessment has been identified as the primary driver of student learning. Paul Ramsden emphatically stated that assessment 'always defines the actual curriculum' so that assessment tasks which require students to demonstrate their command of content through the practical application of legal problem solving skills, for example, are key to the successful promotion of effective student learning.⁵ Accordingly, any assistance that can be provided to explain the expectations of the assessment task will assist the students to effectively and efficiently target their assessment responses. That is, when teachers publish the anticipated standards of assessment criteria necessary to achieve a certain grade – and the student is afforded an opportunity to understand these standards – the student will know what is expected of them and how the teacher will judge their performance. Likewise, Chris Morgan et al have stated that:

[a]s assessment tasks become more centred on learners and their learning needs, students want to know how they are performing against set standards for performance ... Clear standards of

2 During the semester in which this research project operated, the first take-home assignment constituted 40% of the final mark and the final exam constituted 60%.

3 Susan Armstrong et al, *Developing an Assessment Guide to Support the Implementation of a Criteria and Standards-based Approach to Assessment at UWS* (2008) 4.

4 Ibid. See further, D Royce Sadler, 'Specifying and Promulgating Achievement Standards' (1987) 13(2) *Oxford Review of Education* 191, 194; Sadler, above n 1, 160.

5 Paul Ramsden, *Learning to Teach in Higher Education* (2nd ed, 2003) 182.

expected student performance have increasingly become recognised as the vital ingredient in effective assessment.⁶

This view resonates with John Biggs' statement that '[i]n criterion-referenced assessment, students see the situation as a designated learning experience; to get a high grade they have to know the goals and learn how to get there, with a premium on attributions involving effort, study skill and knowing the right procedures.'⁷ Accordingly, the pedagogical reasoning underpinning this criterion-based approach asserts that assessment results will be dependent upon each student learning the appropriate knowledge and skills necessary to satisfy the assessment tasks.

Adherents of criterion-based assessment have commonly asserted that the transparency of assessment criteria and standards to both staff and students alike could be achieved fairly simply through the development and application of explicit assessment criteria and grade descriptors which are tailored to the specific context of a particular unit and discipline.⁸

To this end, a criterion assessment grid – or rubric – was developed in the subject LLB unit that plotted commonly used assessment criteria in matrix format against grades, resulting in grade descriptors that detailed acceptable performance for each criterion at each grade. (In addition to the rubric, various other systems which attempt to map explicit standard parameters include 'grading schemes', 'scoring keys' or 'guides', 'criteria sheets' and 'primary trait analysis').⁹ These grading schemes which employ fixed sets of criteria have become 'firmly established in higher education'.¹⁰ However, the success of this rubric in achieving the effective transmission of meaningful knowledge on assessment standards and criteria to students does not appear to have been an easy nor wholly successful objective. Substantial difficulties were encountered: 'first, in the clear and precise *articulation* of marking criteria and standards and, secondly, in the *accurate receipt* of this understanding by relevant participants undermined their effectiveness of the project (emphasis in original).'¹¹ Berry O'Donovan, Margaret Price and Chris Rust cite their common experience with comprehensive rubric and marking guides and state that they 'failed to transfer meaningful knowledge on standards and criteria' to students (or to individual staff members tasked with marking the respective assignments).¹²

To explain this view, verbal descriptions of standards are always somewhat vague or fuzzy, and often a matter of degree – indicative of relative, rather than absolute positions and are context dependent.¹³ Sadler has argued that fuzzy levels are unavoidable in standards specified in linguistic terms. He states:

[V]erbal descriptions are always to some degree vague or fuzzy ... Fuzzy standards are implied, for example, when it is said that a dissertation is *highly original*, that a solution to a problem is *elegant*, that a student's understanding of entropy is *thorough*, or that a teacher is *competent*. In each of these cases, the standards are designated by linguistic terms ... A fuzzy standard cannot, therefore, be defined into existence. Any attempt to create a formal definition in terms of simpler elements succeeds only to the extent that the simpler elements are easier to understand. However, such a definition has little or no constitutive power and remains essentially tautological because

6 Chris Morgan et al, *The Student Assessment Handbook: New Direction in Traditional and Online Assessment* (2004) 4.

7 John Biggs, *Teaching for Quality Learning at University* (2nd ed, 2003) 59.

8 Armstrong et al, above note 3, 2-3, 6; Jennifer Lingard et al, 'Do Students with Well-aligned Perceptions of Question Difficulty Perform Better?' (2008) *Assessment & Evaluation in Higher Education* 1; Chris Morgan, et al, *The Student Assessment Handbook: New Direction in Traditional and Online Assessment* (2004); Cf O'Donovan, Price and Rust, 'Know What I Mean?', above n 1, 327; Sadler, above n 1, 159.

9 Sadler, above n 1, 159.

10 Ibid.

11 O'Donovan, Price and Rust, 'Know What I Mean?', above n 1, 327.

12 Ibid. See also, O'Donovan, Price and Rust, 'Developing Student Understanding', above n 1.

13 Sadler, 'Specifying and Promulgating Achievement Standards', above n 4, 202; O'Donovan, Price and Rust, 'Know What I Mean?', above n 1, 327.

in the final analysis it rests on interpretations that are based on experience. It is difficult in practice to disentangle the *meanings* of the terms or sentences in a standards specification from their *significance* to assessments of a particular type. To be useful as a standard, assumptions have to be made about the experiences and knowledge background of evaluators and consumers (emphasis in original).¹⁴

Further, Freeman and Lewis have conducted pedagogic research which asserts that the assessment criteria themselves were also subject to multiple interpretations by both individual staff members and students.¹⁵ A response to these difficulties was initially to redraft the grid to more tightly specify the criteria and grade descriptors.¹⁶ However, Sadler has argued that relative terms require an anchor point to communicate definitive standards. In practice, O'Donovan et al found that a 'single-minded concentration on the construction of ever more comprehensive and precise anchor definitions quickly become self-defeating' whereby the very precision of language and terminology 'progressed explication away from common parlance and, as a consequence, definitions became less accessible to novice students'.¹⁷ Secondly, O'Donovan et al continue, 'increasing the quantity of explanation makes for unwieldy and less transferable definitions'.¹⁸

IV. EXEMPLARS AND THE CLARIFICATION OF STANDARDS

Exemplars may be defined as key examples chosen so as to be presented to students as being typical of various designated levels of competency as mapped against the assessment criteria and standards of a given assessment task. Exemplars are not synonymous with standards; rather, they are merely indicative of them – they specify standards implicitly. Put simply, exemplars may be examples of student work which has been previously assessed and then held out to students as being indicative of a definitive assessment standard. Susan Armstrong et al state that:

When formulating standards, exemplars can provide teachers with a starting point to begin articulating different levels of achievement. Exemplars are valuable to students when they are encouraged to consider how the exemplars illustrate the standards described in the marking scheme. Such activities also develop skills in self assessment, providing benchmarks against which students may judge their own performance.¹⁹

On the frontispiece of his book *Teaching for Quality Learning at University*, John Biggs quotes a short passage from Ralph W Tyler which effectively underscores the philosophy of his educative approach and the purpose of his book: 'Learning takes place through the active behaviour of the student: it is what he does that he learns, not what the teacher does'.²⁰ Prosser and Trigwell echo this view by asserting the need for higher education teachers to be more aware of how student learning occurs, and further, to be more aware of the context they create in

14 Sadler, 'Specifying and Promulgating Achievement Standards', above n 4, 202, 204. See also, RE Orton, 'Do Tests Measure Those Constructs that People Interpret Them to Measure?' (1986) 36 *Educational Theory* 233; L Hendrickson 'Procedures and results of an evaluation of writing' (1980) 2(4) *Educational Evaluation and Policy Analysis* 19; Sadler, 'Indeterminacy', above n 4.

15 Richard Freeman and Roger Lewis, *Planning and Implementing Assessment* (1998).

16 Ibid.

17 O'Donovan, Price and Rust, 'Know What I Mean?', above n 1, 327-328; Berry O'Donovan, Margaret Price and Chris Rust, 'The Student Experience of Criterion-referenced Assessment through the Use of a Common Criteria Assessment Grid' (2001) 38 *Innovations in Learning and Teaching International* 74.

18 O'Donovan, Price and Rust, 'Know What I Mean?', above n 1, 328.

19 Armstrong et al, above n 3, 38.

20 Biggs, above n 7, frontispiece. John Biggs quotes Ralph W Tyler, *Basic Principles of Curriculum and Instruction* (1949).

the learning environment.²¹ ‘Preparing a context that will allow students to achieve high quality learning outcomes may necessitate that teachers write and present their teaching in a student focused way’.²²

The use of exemplars – replete with meaningful formative feedback – can facilitate this increased awareness by teachers. Sadler states that students should be educated in the use of exemplars and how to evaluate formative feedback so that they may be enabled ‘to make connections between the feedback and the characteristics of the work they produce, and how they can improve their work in the future’.²³

Similarly, Orsmond, Merry and Reiling assert that the use of exemplars when employing criteria and standards-based assessment can enable students to demonstrate greater understanding of the criteria and standards, thereby achieving higher quality assessment outcomes, to increase their objective ability to engage in self and peer assessment, and to receive more meaningful formative feedback.²⁴

V. PROJECT METHODOLOGY

After the conclusion of the operation of the Contract Law unit in autumn session 2008, the research team requested students from the course to supply the assignments that they had completed for that semester for use in a research project to be conducted in 2009. At this time, students were informed that it was the intention of the research team to use their assignments to construct exemplars for use with future students in this unit and that this process would be documented for research purposes. From over 400 students who completed the Contract Law unit in autumn 2008, the research team received 21 assignments. It was from these 21 assignments that five (representing the grades of fail (F), pass (P), credit (C), distinction (D) and high distinction (HD)) would be chosen for use as exemplars with the Autumn 2009 group of students. Each of these assignments was received in electronic form. The electronic versions of the assignment contained no identification of individual students.

The research team selected the five assignments from the autumn 2008 papers received that they believed best represented the criteria under which marks were awarded for that semester. A detailed process of annotating each of these assignments was then undertaken. This process involved making direct comments onto each of the electronic copies of these assignments which indicated with considerable certainty why the papers received the marks that they did. Both specific and general comments were made onto the papers directly. Both positive and negative specific comments were made on each of the exemplars used.

Specific comments referred directly to individual sentences or parts of the assignment. For example, in the F exemplar, specific comments included: ‘The names of cases and statutes need to be written in italics’ as the student had failed to use an appropriate convention in citing primary sources for their assignment. For the HD assignment used, specific comments included ‘The relevant rationes decidendi are well articulated and contrasted against the significant minority and dissenting judgments.’

By way of general summative comments the F assignment included: ‘This answer fails to sufficiently identify the relevant legal issues and does not apply the law to the facts of the problem scenario. The student does not provide a logical and coherent development of argument and they fail to appropriately incorporate ideas and arguments from both primary and secondary research

21 M Prosser and K Trigwell, *Understanding Learning and Teaching: The Experience in Higher Education* (1999) as cited by Paul Orsmond, Stephen Merry and Kevin Reiling, ‘The Use of Exemplars and Formative Feedback when Using Student Derived Marking Criteria in Peer and Self-assessment’ (2002) 27(4) *Assessment & Evaluation in Higher Education* 309, 320.

22 Orsmond, Merry and Reiling, above n 21, 320.

23 D Royce Sadler, ‘Formative Assessment: Revisiting the Territory’ (1998) 5(1) *Assessment in Education* 77.

24 Orsmond, Merry and Reiling, above n 21, 321.

throughout their entire response. Far too often the student makes a claim and fails to provide any evidence for it. There are also significant problems with referencing of primary sources.’ In this way the summative comment aimed to express in a general sense the overall perception of the assignment with respect to the grade criteria. By way of contrast, the D exemplar included the following summative comments: ‘The student demonstrates considerable knowledge and understanding of the subject matters. Their answer is structured appropriately and displays a coherent argument. They employ a lively and engaging quality of writing’.

In week 5 of the semester, students were asked to complete a survey on their expectations of the value exemplars would provide, if any, to the completion of their assignments. In this week, the seminar leaders engaged each respective class in a discussion of the purpose, form and use of exemplars as an educative tool.²⁵ In week 6, access to the exemplar documentation was made available to all students enrolled in the unit. Access was provided via the electronic Blackboard website for the unit. The documentation they were presented with included a full copy of the 2008 autumn semester assignment and the annotated exemplars for the grades F, P, C, D and HD.

Students submitted their assignment for 2009 in week 9 of the autumn semester. In week 11 they were asked to complete a follow-up survey reporting on their experiences in using the exemplars towards the completion of their assignments. Marked assignments were returned to students following the second survey.

Ethics approval for this study was obtained from the Human Research Ethics Committee.

VI. SURVEY

Two assessment experience questionnaires (AEQ) were used in this study. The first AEQ was utilised to gauge student attitudes to the issue of whether and how they believed the exemplars supported their learning. The AEQ was later re-administered in order to measure the extent of change, if any, in students’ perceptions and responses. Following a diagnosis of the derived data of each of these surveys, it is intended that we may (in our capacity as assessors and teachers) make principled changes to the future assessments in the same unit of study.

Accordingly, the purposes of this study were threefold. The study was designed:

- To investigate the usefulness of exemplars in assisting students to learn the necessary skills to complete their assignment;
- To analyse the usefulness of exemplars in providing students with explicit evidence of the standard of response to a similarly structured assignment within the same academic discipline; and
- To enable students to understand the expectations of the assessment task.

To this end we were interested in evaluating the perceptions that students had of the value of exemplars before they made an attempt on their assignment and then comparing these perceptions with their views after they had formally completed their assignments. In this way our intention was to see the value, from a student’s perspective, of exemplars both before and after being exposed to them.

The surveys used in this study have a basis which draws on the theory of personal constructs formulated by George Kelly²⁶ and further explained by C T Patrick Diamond.²⁷ The surveys are designed to observe the value judgements that students have applied to the questions they

25 Jannette Elwood and Val Klenowski, ‘Creating Communities of Shared Practice: The Challenges of Assessment Use in Learning and Teaching’ (2002) 27(3) *Assessment & Evaluation in Higher Education* 243.

26 George Kelly, *The Psychology of Personal Constructs* (1955).

27 C T Patrick Diamond, ‘Understanding Others: Kellyian Theory, Methodology and Application’ (1982) 6 *International Journal of Intercultural Relations* 395.

are presented with and then to compare these initial values with any change which may have occurred some six weeks later.

At the beginning of the semester 425 students were enrolled for the unit. The precise number of students who were still enrolled during weeks 5 and 11 when the surveys were undertaken is not known; however, the number would almost certainly have been fractionally less than that in the first week of the semester.

Students were presented with 22 items rated on a 5-point Likert scale anchored ‘Strongly disagree’, ‘Disagree’, ‘Neutral’, ‘Agree’ and ‘Strongly agree’, and two open-ended question designed to capture any other comments that students wished to make about their impressions of exemplars at the time. The 22 items were presented to the students in weeks 5 and 11, although the text of the items was slightly altered to acknowledge the timing of the survey (for example the tense of the words used in the questions needed to be altered).²⁸ The use of the open-ended questions was not intended as a method of triangulation or as the basis for a mixed methodology of research finding.²⁹ Instead, it was designed with the intended purpose of gathering any extra information that may have been available at the time. This paper does not examine the information gathered from these open-ended questions.

VII. QUESTIONNAIRE DATA

The dataset consisted of 267 respondents for the pre-assignment survey and 127 for the post-assignment survey. This represents a response rate of 62% and 30% respectively. The following table shows the characteristics of those respondents who completed the survey.

Table 1. Characteristics of Participants who Completed the Pre- and Post-assignment Surveys

	Pre-assignment Survey	Post-assignment Survey
Gender:		
Male	110	53
Female	157	74
Total	267	127
Age:		
18-25	186	98
25-35	62	20
35+	19	9
Total	267	127
Response rate	62%	30%

²⁸ A good example of this occurs with question number 1. In the first survey, undertaken in week 5, the text of the question read “I expect this unit to be difficult”. However, in the second survey, undertaken in week 11, the text of the question read “I found this unit to be difficult.” Clearly the essence of the question is the same, but the phrasing has been altered to reflect the change in tense in the context of the question now being asked towards the end of the semester.

²⁹ Further information on triangulation can be found in Julia Brannen, ‘Combining Qualitative and Quantitative Approaches: An Overview’ in Julia Brannen (ed), *Mixing methods: Qualitative and Quantitative Research* (1992); Todd Jick, ‘Mixing Qualitative and Quantitative Methods: Triangulation in Action’ in John Van Maanen (ed), *Qualitative Methodology* (1983).

All data collected remained anonymous. However, all respondents were asked to indicate their day and month of birth on each of the surveys they completed. This was initially requested in an attempt to collate individual pre- and post-assignment responses and match these. But, given the relatively low response rate to the post-assignment survey, this has not proven practical, nor statistically reliable.

The mean and standard deviation for each of the 22 Likert scale questions posed in the pre- and post-assignment surveys are listed in the following table. Also included is the t-test significance.

Table 2. Results of Pre- and Post-assignment Surveys

Question	Pre-assignment Survey		Post-assignment Survey		T-test significance
	Mean	Standard Deviation	Mean	Standard Deviation	
1. I found this unit to be difficult.	3.41	1.12	3.92	.80	p<0.0001
2. I expect to get a high mark in this unit.	3.53	1.10	3.66	1.12	p>0.05
3. I expect to pass this unit.	4.41	.50	4.78	.20	p<0.0001
4. I know what an annotated exemplar is.	2.54	.40	4.24	.31	p<0.0001
5. Other subjects have used annotated exemplars.	2.41	.43	2.66	.99	p<0.01
6. In the past I have learned what is required for assessment items from the use of annotated examples.	2.29	1.44	3.10	.80	p<0.0001
7. My marks will improve by studying the annotated examples.	3.31	1.12	3.93	.79	p<0.0001
8. Annotated examples have not provided me with any benefit.	2.26	1.42	3.22	.54	p<0.0001
9. I think that annotated examples have helped me to understand what is required in the assessment item.	3.90	.76	4.22	.98	p<0.0001
10. Teacher constructed/provided model examples are the same as examples provided from previous students.	1.51	.45	1.74	.67	p<0.0001

IMPROVING PERFORMANCE IN CONTRACTUAL PROBLEM SOLVING

11. Annotated examples are of no benefit unless they are fully explained by the teacher.	3.53	1.12	2.21	.89	p<0.0001
12. Exemplars increase the level of certainty I have over what is expected of me.	3.51	1.21	4.11	.81	p<0.0001
13. Exemplars provide consistency for all students.	3.72	.32	3.83	.45	p<0.05
14. Exemplars provide me with guidance of how to answer a legal question.	3.84	.65	3.81	.91	p>0.05
15. Exemplars ensure that all students' answers are marked consistently.	2.62	1.33	2.96	.82	p<0.05
16. Exemplars limit my creative ability.	1.95	.67	2.03	.90	p>0.05
17. Exemplars would improve my performance in assignments.	3.31	.65	3.93	.39	p<0.0001
18. Exemplars should not be provided to students until after the assessment item has been marked.	1.67	.43	1.31	.45	p<0.0001
19. I find it difficult to know what markers expect to see in an answer.	4.23	.67	3.73	.99	p<0.0001
20. Exemplars limit my own ability to think.	2.19	1.53	1.30	.98	p<0.0001
21. Availability of exemplars implies more transparency in assessment processes.	3.64	1.12	3.76	1.56	p>0.05
22. I would use any available exemplars as part of my assignment preparation.	3.96	.57	4.21	.41	p<0.0001

The mean and standard deviation of the responses for each of the 22 questions reveals some interesting data. The questions within the survey were broadly designed to cover three main areas. They were designed to judge a student's perception of the difficulty of the unit and their expected results (question set one), their level of knowledge of and previous exposure to exemplars (question set two) and the value they associated with the use of exemplars (question set three). While not necessarily always mutually exclusive these three groupings/sets have been assigned to the following question numbers. Designated as question set one, questions which were designed to record a student's view of the perceived difficulty of the unit and their expected results included numbers 1, 2 and 3. Set two questions which were designed to judge a student's level of knowledge of, and previous exposure to, exemplars consisted of items 4, 5, 6, 10 and 11. Designated as question set three, questions which were designed to explicitly judge a student's individual view of the value of exemplars included at least items 7, 8, 9, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22, although other questions may have intrinsically also indicated a student's view of the value of exemplars.

It would initially appear that very few of these highlighted question sets or individual items have received universal or very strong support from students, with averages for the most part being confined to no more than agree (eg, Question 9 pre-assignment survey) or disagree (eg, Question 6 pre-assignment survey) scores, rather than the maximum possible strongly agree or strongly disagree ranges.

Part of this study was also designed to determine whether students' views towards exemplars changed. Specifically one of the aims of this study was to see if the views of students towards exemplars changed in any sufficient manner from initially being exposed to the exemplar and then completing their assignment. Statistically, as independent samples, the t-test assesses whether the means of two groups are *statistically* different from each other.³⁰ This analysis is appropriate whenever comparing the values of two groups. In this study, students were presented with the same questions pre- and post-assignment completion and the t-test has been used to judge any statistically significant change in their attitudes.³¹

Initially it would appear that many of the survey items reveal significant levels of change (i.e. results where $p < 0.05$) between the pre- and post-assignment results. 18 of the 22 items show a significant to extremely significant change from a t-test significance perspective. 15 of the 22 items show an extremely significant change.

Examining the data in detail, first starting with the results for question set one, averages for the pre-assignment survey consisted of 3.41, 3.53 and 4.41. Averages for the post-assignment survey consisted of 3.92, 3.66 and 4.78. There seems to be clear evidence that most of the students were of the belief that although this may be a difficult unit, they expected to pass. Items 1 and 3 showed an extremely significant change ($p < 0.0001$), whilst item 2 showed no significant change ($p > 0.05$).

The results obtained from question set two are more complex. Item number 4 shows a low mean of 2.54 in the pre-assignment survey, but this increases quite extremely to 4.24 in the post-assignment survey. This results in a statistically significant t-test result ($p < 0.0001$). This may be attributed to the participants actually learning what an annotated exemplar was by being exposed to it in this or other units of study in the time period between the two surveys. However, item 5 would seem to suggest that this increase may be attributed to this unit rather than other

30 For further information regarding t-tests and their statistical significance, see particularly Barbara G Tabachnick and Linda S Fidell, *Using Multivariate Statistics* (3rd ed, 1996) and Thomas Black, *Doing Quantitative Research in the Social Sciences: An Integrated Approach to Research Design, Measurement and Statistics* (1999). Also see Quantitative Methods in Social Sciences, *The T-Test* < http://ccnmtl.columbia.edu/projects/qmss/the_ttest/introduction_2.html > at 13 October 2009.

31 In this study, a two tailed t-test is being employed. For those unfamiliar with the t-test, a p score <0.0001 indicates that a result is extremely statistically significant, a p score <0.01 indicates a result which is very significant, a p score <0.05 indicates a significant result, whilst a p score >0.05 indicates a result which is not considered to be statistically significant.

units as the mean for item 5 increased only from 2.41 to 2.66, suggesting that students had not in fact been greatly exposed to other annotated exemplars during this period. Item 6 results seem to support this further by suggesting that only a relatively small number (at least initially) have been exposed to exemplars in the past.

Item 10 adds weight to these findings as it suggests that students have a good understanding of the different types of exemplars that they could be provided with. Item 10 was phrased, 'Teacher constructed/provided model examples are the same as examples provided from previous students', and the pre-assignment survey mean was 1.51, whilst the post-assignment mean was 1.74.

As question set three contains the largest number of items and directly relates to the purpose of this study, the results obtained need to be discussed in more detail than those of question sets one and two.

In particular, items 7, 9, 17 and 22 relate directly to whether a student believes that an exemplar can assist in improving their result. Pre-survey the means were 3.31, 3.90, 3.31 and 3.96. Post-survey the means were 3.93, 4.22, 3.93 and 4.21. Clearly many students are of the belief that annotated exemplars can help them to improve their results. These results can be contrasted with item 8 which is phrased in a negative manner, but which effectively examines the same concept as items 7 and 9, in order to provide some certainty with the results of items 7 and 9. Item 8 scores a mean of 2.26 in the pre-assignment survey, but that mean did increase to 3.22 in the post-assignment survey. The scores obtained from these five different, but correlated, items add considerable weight to the value of the data which can be gathered from any one of the items.

An argument might be made that exemplars may be too prescriptive and limit a student's own ability to think or be creative, or even to answer a problem in a novel or a unique fashion. Items 16 and 20, with mean scores of 1.95 and 2.19 in the pre-assignment survey and 2.03 and 1.30 in the post-assignment survey, seem to indicate that students do not believe that exemplars limit their creative ability. These results can be contrasted with items 14, and 19, which were centred on a student's inability to understand what a marker may be looking for. With mean scores of 3.84 and 4.23 for the pre-assignment survey and 3.81 and 3.73 for the post-assignment survey, this would indicate that even if students were of the view that exemplars limited their creative ability, they may be prepared to sacrifice this for the benefit of actually knowing what it was that was needed to gain them high marks in an assignment.

Items 13 and 21 were designed to gauge the student's perceptions of the value of exemplars from the perspective of ensuring that their work would be treated equally and marked fairly in comparison to expected norms. With mean scores pre- and post-assignment survey of 3.72, 3.64, 3.83 and 3.76, there appears fairly clear evidence that students were of the belief that exemplars could be valuable for this purpose.

Finally, item 8 provides an interesting contrast to all of those other items discussed above. This item was designed to measure overall the student's view of the inherent value of annotated exemplars. The item was phrased, 'annotated exemplars have not provided me with any benefit'. The pre-assignment mean was 2.26 and the post-assignment mean 3.22. Overall this could hardly be considered to be an outstanding endorsement for the use of annotated exemplars.

While the mean scores obtained from each of the relevant items for this question set are important and do present valuable data in themselves, some of the more pertinent and valuable data is gathered by looking at the t-test significant changes measured in this question set. Fourteen questions make question set three. Of those 14, only 3 do not register at least a significant change in value. So the result is that 11 of the 14 do register at least a significant change. Quite importantly, 9 of the 14 actually registered an extremely significant change. This is a very large proportion and can be used to demonstrate some significant changes in students' perceptions towards the value of exemplars after being exposed to the exemplars and then completing their assignments.

Within question set three, items 7, 8, 9, 12, 17, 18, 19, 20 and 22 registered an extremely significant change. Primarily, this demonstrates that the views of the participants about the value they associated with the use of exemplars underwent significant change. This significant change may be associated with their association to the annotated exemplars used in this research project or to some other external factors. Nonetheless, extremely statistically significant changes in attitudes have occurred. However, the authors are conscious of drawing attention to the relatively low response rate, at least compared to the original survey, for survey completions.

VIII. DISCUSSION AND CONCLUSION

This research paper has reported on the value of using annotated exemplars for students completing a law degree. The literature discussed in this paper suggests that students can benefit from being exposed to exemplars during a course of study. Specifically, it appears from the literature that exemplars can enable or assist students to demonstrate greater understanding of the assessment criteria and standards, aid students to achieve improved academic performances, provide students with increasingly more meaningful formative feedback on assessment performance, and facilitate students engaging in critical and objective introspective and peer judgment of assessment performance.

However, the data from this research paper is not as conclusive. There does not appear, from a student's perspective, to be overwhelming demand for, or understanding of, the use of annotated exemplars. It may be that students do not fully understand the value of annotated exemplars or even understand the differences between teacher-constructed 'model' answers, exemplars and annotated exemplars and may need to be educated about these things in order to understand the intrinsic value associated with annotated exemplars.

Although the data from this research paper does not demonstrate overwhelming support for annotated exemplars, what is perhaps most interesting is the significant changes in the responses recorded from the pre- and post-assignment surveys. With 18 of the 22 items recording at least a significant change and with 15 of 22 recording an extremely significant change, this may suggest that although students do not display an overwhelming or universal support for annotated exemplars, they do demonstrate a change in attitude after being exposed to them. Perhaps as students become increasingly exposed to exemplars and see the different aspects associated with them, their attitudes will change again. Students may also develop the ability to be able to discriminate between good and bad exemplars.

Specific attention is drawn to the fact that the post-assignment survey was conducted before students received their assignments and knew of their marks for this assessment activity or for the unit as a whole. Survey results may have been significantly different if the survey was conducted post this event. As researchers we are keen to examine whether the annotated exemplars used in this unit during this semester had an actual effect on the marks that students obtained. To this end, a future research project will examine this aspect.

Annotated exemplars take a considerable amount of time and effort to construct. Specific exemplars need to be carefully chosen from the available samples, student consent to use these as exemplars needs to be obtained, and the exemplars need to be annotated, placed into a presentable format and distributed to students. The results from this research paper do not clearly demonstrate that students find this product of fundamental value.