

Does providing exemplars to students lead to a change in their grades? An examination of the available data on this question

Dr. David Newlyn

School of Law, Western Sydney University, Sydney, Australia

Abstract

Exemplars have been in existence for some time. That students want and value exemplars has not been a matter of significant contention. But whether exemplars actually lead to any change in the marks or grades that students achieve is a much more problematic issue. There have been very few well-constructed, thorough, empirical studies into this question. This article analyses three recent studies that have been undertaken into the value of exemplars in leading to changes to students marks, discusses the results of these studies, their limitations and concludes with some comments about how this question can be answered.

Keywords: exemplars, feed-forward, assessment, higher education, evaluation, marks, grades

1. Introduction

'Unless students have a concept of what good work looks like, it is difficult for them to produce quality assignments'^[1] This statement provides a somewhat felicitous starting point for this article as the value of using exemplars to assist students to learn, across multiple forms of assessment, has already been well established.^[2, 3] As Newlyn indicates, exemplars are '...specific examples (of work) chosen 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.'^[4] That exemplars are a very powerful means by which to impart students with particular academic standards, criteria and skills is not in doubt.^[5, 6, 7] But whether these same exemplars can lead to a change in student marks is a separate matter which has received little attention.

This article reports on three different studies which have been undertaken in an effort to determine the value of exemplars in leading to changes (positive or negative) in the marks awarded to students. This article provides a detailed report and critique of the limitations of these studies. The article ends with a discussion of the current state of whether exemplars can lead to a change in students' marks as well as a discussion of how we may in the future be able to more accurately determine an answer to the question of determining the explicit value of using exemplars.

2. Wimshurst and Manning study

In a 2013 article published in the *Assessment & Evaluation of Higher Education* journal, Wimshurst and Manning described a study that they undertook that commenced in 2010 with a group undergraduate students undertaking an elective law unit of study at an Australian university.^[8] That study involved the provision of six unannotated exemplars to students which ranged from a mere pass to the top grade level available. No example of a fail grade was made available to students. The exemplars used in their study were constructed from previous student's submitted work. The exemplars were made available to all students enrolled in the unit via the course website.

The authors of the study, inter alia, sought to measure the change in student's grades as a result of being provided with the exemplars. To ensure that all students engaged with the exemplars provided, Wimshurst and Manning made using the exemplars a part of an assessment activity. That is, all 97 students enrolled in the unit had to use some of the exemplars as part of an assessment activity for the unit that they were enrolled in. Specifically, it appears that students had to select three of the exemplars and 'mark' those exemplars as if they were assessing the grade as a tutor may in a real world environment. Along with providing a short (100 – 150 words per exemplar) explanation for their justification of the mark awarded, students also had to provide each exemplar with an individual mark. Several weeks' later students were provided with feedback on their assessment task. That feedback showed the range of marks that other students had awarded to the same exemplar and the marks that staff had originally awarded to the same exemplars. In addition some comments derived primarily from the student's observation of these exemplars were also provided as part of the feedback. Students then used this feedback as a basis for completing a further element of assessment.

Before examining the impact of the exemplars, several things should be noted about the Wimshurst and Manning study. The first is that at no stage were students actually provided with annotated exemplars. The second is that students did not see the grades awarded for the exemplars provided to them until after they had completed their first assessment task which was based on the exemplars provided. The third is that students were asked only to work with three of the six available exemplars and these three exemplars may not have represented the full range of grades available. These are important points of distinction which should be born in mind when considering the results obtained from this study and contrasted to the different ways in which exemplars were used in other studies reported in this same article.

In order to gauge any change in student grades, Wimshurst and Manning compared the results of the 97 students who were provided with exemplars in 2010 against those 87 students

who completed the same unit in 2009 who were not provided with exemplars. After attempting to compare the statistical similarities between the 2009 and 2010 groups, it appears that there is some evidence that providing exemplars to the 2010 group helped to improve marks. However, apart from an attempt to link different changes in marks to a student's GPA, the authors of the study do not provide sufficient detail to accurately measure the significance of this change across the entire cohort of students or for individual students, except to indicate that grades changes by 7%. An examination of the number of different grades awarded between the two years of the student and/or an examination of the means/mode or other relevant mathematical measure may have been more useful in determining the success of using exemplars. In short, the authors of this study provide limited evidence for their claims that the use of exemplars lead to a positive change in the marks obtained by students

The authors conclude their article with the claim that they aimed to measure the value of using exemplars in a 'rigorous' fashion. However, even given the use of qualitative data in the study (not reported here), noting the limitations indicated above, it is unlikely that this claim can be adequately sustained.

3. Kean study

In a 2012 article, Kean reports on her experience using exemplars with a group of journalism students at a university located in the United Kingdom.^[9] Kean had been influenced by the works of Sadler^[10, 11, 12] and set out to measure the impact of exemplars by examining the views of students and by measuring the impact of exemplars on the marks awarded to students.

As part of a gradual process of formative assessment, Kean aimed to increase the students' knowledge of content and skills by gradually introducing them to exemplars which would be used to complete the one item of assessment worth 100% of the marks available for the unit of study. This assessment item was an unseen examination.

The 41 students enrolled in the journalism unit undertook a mock exam. Their exams were returned to them, but no marks were provided. Instead what was provided were three exemplars which were constructed from the answers students who had undertaken the mock exam in the previous year had submitted. The three exemplars represented results of 68%, 52% and 35%, although these results were not known to the students at this stage. Students were asked to grade those exemplars using marking criteria provided to them. They were then asked to grade their own paper using the same criteria. At the end of the semester of study when students completed the final examination the average was 56%. In the previous offering of the same unit when exemplars had not been used the average was 50.8%. Kean also reported that the failure rate had improved when exemplars had been provided. When exemplars were provided the failure rate was 5%. When exemplars were not provided the failure rate was 15%.

The Kean study represents a rather limited example of how the value of exemplars may be measured in terms of impacting upon students results. Firstly it should be noted that the study involved only 41 students. Secondly, Kean misses opportunities to provide a though discussion of the methodology used. Important details such as why only three exemplars were provided to students are missed as well as

exactly how the exemplars were constructed and issued to students. It is these types of specific details which would make replicating the experiences of Kean possible. Thirdly it should be noted that a rather simplistic analysis of the impact on results has been undertaken by Kean.

This is exaggerated by the fact that, despite acknowledging the importance of it, Kean spends just three lines out of an eleven page article discussing the impact of exemplars on student's results. Clearly Kean was interested in measuring the impact, but she appears to have lacked the understanding of how this could be undertaken in a meticulous or punctilious approach. Kean provides details of only how using exemplars impacts on the average and fail rates of students. Other details such as the change in different grade levels could also have been valuable. That is, did more students access the top grade levels in the year in which exemplars were used? Disappointingly, Kean has missed a significant opportunity to rigorously examine the impact of exemplars. Whilst it could not be said that the Kean study represents anecdotal evidence of the impact of exemplars on student results, it could be argued that the Kean study represents limited evidence of the impact because of the restricted way in which Kean worked with the available data.

4. Scoles, Huxham and McArthur study

In 2013 Scoles, Huxham and McArthur reported on their analysis of exemplars used in a higher education unit which was conducted at a Scottish university.^[13] Exemplars were introduced into a life sciences unit in order to examine the benefit they had on the results achieved by students specifically in relation to their performance in the final exam. The authors of this study provided exemplars to student before the final examination was held. The exemplars were provided in an online Web CT environment and the downloading of the exemplars was recorded. The study involved 520 students.

The exemplars used in this study were constructed from the exam papers of previous students. After obtaining student consent and de-identifying the answers, three exam papers were used as the exemplars provided to the students. These represented an average (50 to 60%), good (60 to 70%) and an excellent (70% and above) response. No example of a fail was provided to students, this appears to have been because the authors of the study believed it may be difficult to receive the consent of a student to provide such an example. Notably, the exam answers had been transcribed and typed into a new document in order to prevent any difficulties which may have arisen from students being unable to read the handwriting of the original exam answers submitted by students. Further the exemplars provided to students were annotated to reflect the original comments provided by the examiner when the original exam papers were marked. In order to measure the impact of exemplars on student marks the results of those students who accessed the exemplars were compared to the results of those students who did not access the provided exemplars.

The authors of this study noted that 76% of students accessed the provided exemplars. Reporting on how the exemplars impacted on student results it was noted that the mean mark of those who accessed the exemplars was 54.8. Students who did not access the provided exemplars achieved a mean of 48.7. The standard deviation for those that accessed exemplars was 15.2, whereas the standard deviation of those that did not access the exemplars was 20.3. The range for those accessing

the exemplars was 93. Whereas the range for those not accessing was 89.

A number of things should be noted from the study undertaken by Scoles, Huxham and McArthur. The first is that of all of the studies reported in this article, this is the only study which provided annotated exemplars. Annotated exemplars are those exemplars which actively seek to demonstrate to students why a particular mark or grade was awarded. Non annotated exemplars are simply examples. That is, they do not seek to explain but rather just demonstrate and hope that students, almost via a process of osmosis, understand what is being shown to them. Secondly, this study does not attempt to compare the results of students with those in a previous cohort. Rather a comparison is made with students in the same year who did not access the provided exemplars. The problem with this approach is that it cannot be said for certain that these two groups are mutually exclusive. Take for example a situation where one student downloaded the exemplar document. This student would therefore belong in the group assigned as having accessed the exemplar. That student could have provided this exemplar to another student directly. That other student would not be recorded as having accessed an exemplar. Thirdly, although a larger population sample is provided for in this study, the range of data reported is limited. The authors of the study report on only the mean, standard deviation and range achieved. An examination of other indicators, such as mode, median or even differences in grade levels achieved (high distinction, distinction, credit, pass, fail or similar) may also have been very useful in assisting to demonstrate any change in student results.

5. Common problems with research into the value of exemplars

The reality is that determining the value of using exemplars is a complex problem. The three studies reported in this paper demonstrate that there can be a variety of different methods in which to investigate the question of whether providing exemplars to students' leads to a change in the marks that they receive. Each of the three studies reported have used different methods which seemed appropriate to the authors of those studies. And each of those three studies reports appropriate results for the methods chosen. However each of the three studies has its limitations. Legitimate questions about the authenticity of the results of these three studies can be raised. It is likely that if these same three studies were to be replicated different results would be achieved.

It may never be possible to provide a definitive answer to the question of whether providing exemplars to students leads to a change in marks as there are a large number of variables to be considered and many of these variables may be insurmountable. In order to determine the value of the results reported in the studies in this paper questions such as the following should be asked:

1. Was the unit an elective or a core unit?
2. How many students were enrolled in the unit?
3. How many exemplars were provided to the students?
4. How were the exemplars constructed – teacher or student constructed exemplars?
5. How do you measure who 'looked at' the exemplars?
6. How do you measure whether students 'used' the exemplars?

7. How many years of data do you need in the same subject area to ensure the reliability of the data?
8. Should exemplars in different subject areas be measured/used in different ways?
9. How do you measure the changes in results? Should you measure the change to grade levels awarded or to averages/median/mode?
10. Is the background of the student important? Are issues of gender, race, religion, or entry scores to the unit relevant?
11. Would the student have achieved the same result if exemplars were not provided? How would you determine this?
12. What if you provided exemplars to some students in a course but not to others? Is this a valid method of measuring the impact of exemplars? What are the ethical implications of knowingly providing inferior learning resources to some student?
13. How many students are needed to constitute a valid study? Does the number need to be equal each year?
14. Is it relevant that the person who marked the assessment item is a partner in the study of exemplars?
15. Do the people who construct the exemplars have a 'vested' interest in the success of the exemplars?

Clearly the above list is not exhaustive. No doubt other equally important questions about the limitations of the data presented could also be raised. This then may lead to the question "What is the perfect way to determine whether or not providing exemplars to students leads to a change in the marks awarded?" To this question there may be no 'perfect' answer as the number of variables involved and the criticisms that could be made of the methodologies chosen for investigation are infinite. The most that may be hoped for is the 'best' or 'most reliable way'. To this end it is suggested that whilst studies like those reported in this article are important, what is perhaps more important is that they be replicated. That is the data that has been obtained using the specific methodologies listed in each of these studies needs to be replicated with other groups and by other researchers. There may be limited value in attempting to design the 'perfect' basis for studying the value of exemplars as enough 'different' ways in which this can be achieved has already occurred. Instead what would be more beneficial is that others replicate the methodologies already used in their own fields of study. It may be only if this occurs that we will be better able to determine the impact of exemplars on student's marks. Any study into the impact of using exemplars will always have limitations. If those limitations are minimised and noted, then the data obtained can be interpreted in this context.

6. Conclusion

There should be no doubt that exemplars are wonderful things. Exemplars have significant value. Exemplars convey information in a manner in which nothing else can. There is enough evidence available already to suggest that they are valued by students. Through their myriad of different types they can communicate to students in a variety of ways, imparting knowledge and skills in a special manner. They provide a rich source of immersion into an unknown environment for students, thrusting them into a world of clear examples of best practices which are designed to provide clarity, precision and certainty as to the skills and competencies expected of them.

Analysing the effectiveness of exemplars in the classroom is a difficult task, as learning takes place in a complex nonpareil environment. There are a multitude of factors which can influence the impact that exemplars have on students and then flow through to measuring any change on a student's marks. This article has reported on three studies undertaken in a number of different fields where attempts have been made to quantify the value of using exemplars on the results students obtain. Each of the studies reported on have very different methodologies. Whilst it was not the intention of this article to specifically attack any or all of the reported studies, each of these studies is open to legitimate criticisms because of the methodologies chosen. None of the studies reported on could be described as 'perfect'. That is they all have limitations. But it is also promulgated in this article that there may be no opportunity for a 'perfect' study that could be undertaken in order to ascertain an unquestionable answer to the question of whether providing exemplars leads to a change in student's marks. Instead what has been suggested in this article is that the methodologies used in the studies be replicated in other environments and via an examination of the increased range of data that would be generated in these circumstances a more accurate view of the impact of exemplars would emerge.

At present, there remains some relevant evidence that the use of exemplars can lead to an impact on the marks achieved by students. However, this evidence is limited and needs further examination in a rigorous systematic manner.

7. References

1. Carless D, Chan KKH. Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*. 2016; 0(0):1-12. <http://dx.doi.org/10.1080/02602938.2016.1211246>.
2. Handley K, Williams L. From Copying to Learning: Using Exemplars to Engage Students with Assessment Criteria and Feedback. *Assessment and Evaluation in Higher Education*. 2009; 36(1):95-108.
3. Newlyn D, Spencer L. Using exemplars in an interdisciplinary law unit: listening to the students' voices. *Journal of the Australasian Law Teachers Association*. 2009; 2(1):121-133.
4. Newlyn D. Using Exemplars in a University Environment: The Good, the Bad, and the Ugly. *Journal of the World Universities Forum*. 2014; 6(3):39-49.
5. Huxham M. Fast and Effective Feedback: Are Model Answers the Answer? *Assessment and Evaluation in Higher Education*. 2007; 32(6):601-611.
6. Donovan OB, Price M, Rust C. Know What I Mean? Enhancing Student Understanding of Assessment Standards and Criteria. *Teaching in Higher Education*. 2004; 9(3):325-335.
7. Newlyn D. Is there a correlation between providing exemplars and a change in student marks? A report on a longitudinal study designed to answer this question. *International Journal of Research in Humanities, Arts and Literature*. 2017; 5(2):73-80.
8. Wimshurst K, Manning M. Feed-forward assessment, exemplars and peer marking: evidence of efficacy. *Assessment and Evaluation in Higher Education*. 2013; 38(4):451-465.
9. Kean J. Show AND tell: using peer assessment and exemplars to help students understand quality in assessment. *Practitioner Research in Higher Education*. 2012; 6(2):83-94.
10. Sadler DR. Interpretations of Criteria-based Assessment and Grading in Higher Education. *Assessment and Evaluation in Higher Education*. 2005; 30(2):175-198.
11. Sadler DR. Formative Assessment: Revisiting the Territory. *Assessment in Education*. 1998; 5(1):77-84.
12. Sadler DR. Grade Integrity and the Representation of Academic Achievement. *Studies in Higher Education*. 2009; 34(7):807-826.
13. Scoles J, Huxham M, McArthur J. No longer exempt from good practice: using exemplars to close the feedback gap for exams. *Assessment & Evaluation in Higher Education*. 2013; 38(6):631-645.