

How to Formulate Relevant and Assessable Learning Outcomes in Statistics

Thomas E. Holgersson

Department of Economics and Statistics, Linnaeus University, Växjö, Sweden
Email: thomas.holgersson@lnu.se

Received 4 March 2016; accepted 19 April 2016; published 22 April 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Course syllabuses, outlines or similar academic documents specifying the content of a course will often be a helpful tool both for teachers and students to grasp the content and purpose of a course. In many cases, however, the compilation of such documents is a painstaking process for the educator designing it, and is a task that many teachers will shun. In this paper we propose a fairly simple pedagogical model for designing specific learning outcomes that the students are expected to attain after completion of a course.

Keywords

Learning Assessment, Course Syllabus, Statistics

1. Introduction

The academic world has long worked with evaluation and quality issues in various guises. A large part of this task falls on the individual course coordinators, which are often the same as the teachers. This, in turn, means that the pedagogical tasks are undertaking a shift from teaching towards issues of evaluation and Teachers are faced with the compilation of course syllabuses or course outlines, which in many cases is the only formal agreement between the individual student and the university, at least on the course level, and therefore very important. A specific matter concerns the formulation of learning outcomes, since these constitute a direct link between course content and the examination. Clearly defined learning outcomes can provide a strong support for both the teacher and the students. They also have a direct bearing on the examination: are the right things being assessed, and to what extent is the right things assessed? An old saying puts its very eloquently: “If you don’t know where you’re going, any bus will do.” This can also be applied to the planning of a teaching programme or a specific course. Baume (2002) stress that “If a lecturer doesn’t know what he or she is trying to achieve, then almost any module structure, or any teaching method, will do”. In practice, it is not an easy task to formulate learning outcomes that

are both assessable and relevant. This article refers to as a contribution to this issue, with the delimitation to the topic of statistics at the introductory level. The reason for this is that the content and format of statistics courses quickly changes from emphasis on general understanding and applications, to become more mathematical in nature at a higher level (“theorem-proof” problems).

2. Models for Developing and Assessment of Learning Outcomes

The problem of assessing and evaluating the quality of education and learning has been dealt with in many different contexts. In the literature there are many rather different traditions of research on learning processes. In some cases focus has been set on strategies for surface and deep learning (e.g. Marton & Säljö, 1976a,b; Biggs 1987, 1993) and in other cases on hierarchical learning levels, or domains, described through taxonomies where each level is a subset of the previous one. The first systematic way of describing how learners performance develops was probably the taxonomy developed by Bloom (1956), who sought to create a more holistic view on education. His taxonomy has been extensively used in teacher education to develop learning and teaching strategies. Blooms taxonomy was further on developed and redefined by Anderson and Kratwohl (2001) who added ideas on how the taxonomy intersects and acts upon different types and levels of knowledge.

A somewhat different way of describing how learners performance develops was introduced by the SOLO (Structure of the Observed Learning Outcomes) taxonomy of Biggs & Collins (1982), which is a theory about teaching and learning rather than a theory about knowledge. It describes a hierarchy which has 5 stages or levels for the purpose of assessing the students learning based on the quality of their work. The SOLO taxonomy is based essentially on various types of systematizations and constructive alignment of verbs that can be used for learning, teaching and assessment of knowledge. This is done in four steps: (i) defining the intended learning outcomes in the form of standards that students are expected to achieve (usually expressed in terms of appropriate verbs) and (ii) creating a learning environment and appropriate activities that is likely to lead to the intended learning outcomes (iii) assessing the extent to which students actual learning outcomes match the intended learning outcomes, and (iv) rating the student’s achievement. The taxonomies by Bloom (1956), Anderson & Kratwohl (2001) and Biggs & Collins (2007) discussed above are all intended to be applicable in any scientific discipline. But this generality may come at the cost of not being completely adapted to any discipline. This fact has been stressed by Bloom, who argued that “*Ideally each major field should have its own taxonomy in its own language-more detailed, closer to the special language and thinking of its experts, reflecting its own appropriate sub-divisions and levels of education, with possible new categories, combinations of categories and omitting categories as appropriate*” (Kratwohl & Anderson, 2001). In Statistics, for example, it frequently happens that students can easily memorize a technique, say for calculation of a confidence interval, and then reproduce it without having a slightest idea of what he/she just done. In this view, “Apply” and “Understand” in the Anderson & Kratwohl taxonomy in Figure 1 below should swap places. Other taxonomies and strategies that are more adapted to problem-basted subjects have also been proposed. One example is the Feisel-Schmidt taxonomy by Crawley (2007), but not even this is completely fit for Statistics. It should be obvious that no strategy is completely applicable to all subjects, a fact that has been stressed by Elmgren & Henriksson (2010). At the same time, each of the above mentioned taxonomies, and in particular the SOLO taxonomy, could well be used as an important component in a strategy for constructing assessable and relevant Intended learning outcomes (ILO) in Statistics. In this paper, however, we will set up a more hands-on strategy, rather than a taxonomy, for designing assessable and relevant Intended learning outcomes in Statistics that should be simple and accessible enough to be implemented by the typical course coordinator or lecturer in Statistics.

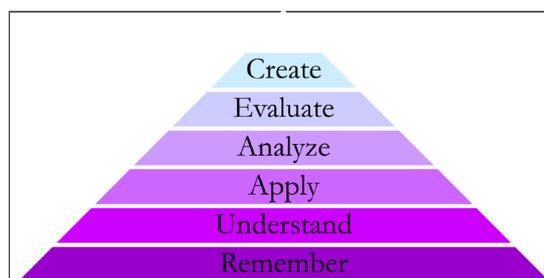


Figure 1. Visual taxonomy by Anderson and Kratwohl (2001).

The issue of how to formulate relevant and assessable learning outcomes in Statistics obviously consists of two parts, *relevant* and *assessable*, and therefore our strategy will take its standpoint in these two. The fact that an intended learning outcome is relevant does not imply that it is assessable and vice versa. The issue is made further complex by the fact that the learning outcomes within a course cannot be separated from the learning process. For example, cleverly defined intended learning outcomes should permit the individual student to choose from several types of expressions. The student may explain a certain concept either verbally *or* mathematically/symbolically. This, in turn, presumes that the ILO's are designed to permit such options. Although the matters of learning process and adaption to the students' individuality partly lie beyond the matters of *relevance* and *assess ability*, they must always be kept in mind while designing the ILO's.

Relevance. In order for the ILO's to be relevant they must first and foremost involve knowledge, skills and approaches connected to the contents of the course. Hence, in order to ensure the relevance of the ILO's, the course contents should be listed in one way or another. On a schematic plan, the general course content is determined by the programme manager, course coordinator, external mandator of the course or similar. But it is also important that the content is specified at a detailed level to facilitate formulation of assessable learning outcomes and provide students with a better understanding of what is being expected of them. The course literature is a very useful tool for this purpose. The Content index, which is normally available in the very beginning of the course book is a convenient source, that may after a careful selection of relevant parts be incorporated in the course syllabus (or course outline in academic environments without a tradition of using formal course syllabuses). A typical book in introductory level Statistics (e.g., Aczel & Sounderpandian, 2006; Bland, 2000; Howitt & Cramer, 2010) usually consists of 10 - 20 chapters, each consisting of 5 - 10 sub-headings. These provide a large number of lines that can be advantageously used as a source in the formulation of the course contents. This procedure of firstly taking into account the general objectives of the course, then deciding the literature and finally setting the detailed contents in the course through a course syllabus or similar document, facilitates in many respects the development of the course, and also ensures that the literature becomes more integrated into the course, rather than being only a peripheral supplement to the lectures. By combining the headlines with suitable verbs one can in a relatively simple way ensure that the ILO's are tightly linked to the course content. It is also important to remember that students have a right to know what is expected of them. A strong connection between the literature and the intended learning outcomes may substantially help the student to comprehend and view the ILO's. At this point it should be emphasized that the course content, of course, cannot itself constitute the ILO's. They need to be problematized by the lecturer, put into the right context and properly linked to each other. The point made here is that the literature should play a central role when determining the contents of the course and thereby also be a direct link to the intended learning outcomes. This approach should be contrasted with the procedure in which the course content is presented only briefly by the teacher and irrespective of the literature (sometimes subject to the teachers improvisation) which leaves the student in a state of uncertainty. Such a procedure also risk making both teaching and assessment strongly dependent on an individual teacher, and gives the student a rather vague idea of what is being expected of them.

Assess ability. Because an assessment task, such as a question in a written exam is always linked to verbs in one way or another they deserve some attention in its own right. The website of the [American Association of Law Libraries](http://www.allnet.org), (<http://www.allnet.org>) argue that "*Since the learner's performance should be observable and measurable, the verb chosen for each outcome statement should be an action verb which results in overt behavior that can be observed and measured. Certain verbs are unclear and subject to different interpretations in terms of what action they are specifying. ... These types of verbs should be avoided*".

After completing a course the student who attains the learning outcomes and pass exam is expected to possess certain skills, knowledge and abilities and must therefore prove these in an examination task. Expected learning outcomes may therefore be expressed as, e.g., that the student after completion of the course should have *knowledge of*, or *be able to* accomplish something. It is obvious that verbs such as *know* or *be able to* are extremely difficult to assess. How can a teacher who is grading an exam know whether or not a student *knows* something at a sufficient level? These are examples of so-called "vague" verbs because they do not allow for adequate assessment. The opposite to vague verbs are "strong" verbs, such as e.g., *compare* or *calculate*. Hence it comes natural that choices and uses of different verbs has become a central ingredient in the research of learning, teaching and assessment. Examples include the taxonomies by Bloom (1956), Anderson & Kratwohl (2001) and Biggs & Collins (1982), where the different levels are demonstrated and exemplified through verbs. Biggs & Tang (2007) list some specific vague verbs, such as *Familiarise with*, *Appreciate*, *Become aware of*, *Know*,

Learn about and *Understand*. These verbs should be avoided when specifying learning outcomes. At the same time, complete absence of vague verbs in ILO's may actually shift the teaching towards surface learning, and thereby contradicting what one originally meant to achieve. Compare the following two sentences: (i) *On completion of the course, the student is expected to know the meaning of a probability* vs (ii) *On completion of the course, the student is expected to be able to calculate elementary probabilities*. The second ILO centers about a strong active verb (*calculate*) while the active verb in the first sentence is vague (*know*). By using the second intended learning outcome there is a risk that a student may pass the question without even being able to explain what is meant by a probability. This at the same time as the verb *calculate* is much stronger than *know*. In fact, strong verbs are frequently highly assessable and easy to grade, but may at the same time drive the students towards surface learning. When an individual teacher faces the task of formulating ILO's then (s)he probably strives towards ILO's that are not only assessable but also stimulate students towards deep learning. In this view it makes sense to classify verbs not only as being strong or vague, but also to the level of which they drive the students to deep learning. Here we define verbs that stimulate deep-learning as "deep verbs" (the antonym being surface). **Figure 2** below demonstrates this. The x-axis moves from vague verbs on the left towards strong verbs on the right. The y-axis moves downwards to more deep verbs. The left uppermost quadrant of **Figure 2** consists of vague verbs that are not strong. Verbs belonging to this class will usually have little relevance in terms of ILO's. Here we exemplify that quadrant by *know* and *realize*. These verbs cannot be said to stimulate deep learning nor can they be easily assessed. Moving on to the right uppermost quadrant the verbs are strong but not deep. Examples include *calculate* and *conduct*. ILO's involving these verbs would typically be easy to assess but cannot be said to stimulate deep learning. The left lowermost quadrant represent deep but vague verbs, here exemplified by *reflect*. If students are instructed to *reflect* about a concept, such as frequentistic vs Bayesian view, they would likely problematize the concepts at a fairly advanced level, but the ILO involving *reflect* might turn out difficult to assess and grade. Finally, the lowermost left quadrant contains strong and deep verbs. These are here referred to as "dynamic" verbs. Examples include *explain* and *contrast*. Students asked to *compare and contrast*, e.g., parametric vs. non-parametric methods, are not likely to pass the question by plain memorization of words and phrases. At the same time, *Explain, compare and contrast parametric vs. non parametric methods* should be relatively easy to assess in a meaningful way.

At this point it should be stressed that the concepts of "deep" and "strong" verbs are highly contextual. To *be able to find* a minimal sufficient statistic (Cox & Hinkley, 1973) is both deep and strong while to *be able to find* a critical value from a *t*-table must be considered as strong but certainly not deep (it would be very simple to memorize the procedure of reading the table without having a clue of its meaning or use). The classification of verbs on the "deepness axis" is surely more contextual than classification on the "strongness axis". Nether the less, given a certain subject to assess, **Figure 2** should be a helpful tool to classify verbs. When setting a number of ILO's to a statistics course one should avoid having too many verbs in the left uppermost quadrant but include at least some dynamic verbs (belonging to the right lowermost quadrant).

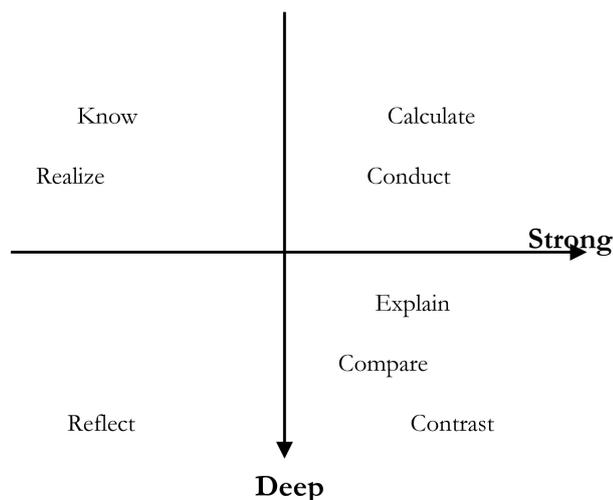


Figure 2. Strong and deep verbs.

Formulation of relevant and assessable intended learning outcomes in Statistics. The discussions in the sections above are here compiled to a simple model for formulation of relevant and assessable intended learning outcomes, the criteria's being that the ILO's should be strongly linked to the content of the course, should facilitate teaching styles that stimulate the students to deep learning and also be assessable in a meaningful way. The first step takes its standpoint from a specific concept from the contents of the course. Once the main content of the course has been set the ILO's should be coupled to the content. It will be convenient to set up about 2 - 4 general questions of the kind "When", "Why", "What" etc. The learning outcomes should then be associated with at least one such class. An example of the procedure is given below.

Example: Intended learning outcomes concerning Non-parametric methods

Suppose that an introductory level course in Statistics includes a module of Non-parametric methods for inference of measures of centrality, and that learning outcomes hence should be associated with this subject. The methods within this family are characterized by the facts that they involve a minimum of distributional assumptions, they usually have exact test level regardless of the sample size but have high type-II error probability and should hence not be used habitually. One objective of the course is therefore that students should have insight into *when* nonparametric methods should be used and *why* they are used. And also *how* they are used and *what is meant* by non-parametric methods. This is visualized in **Figure 3**.

Some intended learning outcomes associated with these questions becomes:

On completion of the course, the student is expected to be able to:

- Compare and contrast parametric and non-parametric methods;
- Explain when and why non-parametric methods should be used;
- Give an example of a situation when non-parametric methods (not) should be used;
- Explain how non-parametric methods should be used;
- Conduct a non-parametric test;
- Give an example on non-parametric methods for hypothesis tests'

Note also that the above ILO's easily can be re-expressed into exam questions, for example:

Compare and contrast the sign test and *t*-test;

Explain when and why non-parametric methods (not) should be used;

Explain how the Sign test, Wilcoxon signed rank test and Mann-Witney's U-test are conducted;

Given the context *X* and the data set *Y*, use the sign test to test the null hypothesis that the median $M = 0$;

Give examples of two non-parametric methods for testing hypotheses about the median, and state the necessary distributional assumptions.

The above exam questions need, of course, to be further clarified and contextualized, but never the less show how a learning outcome related examination may be outlined. The advantages of the approach are that it provides

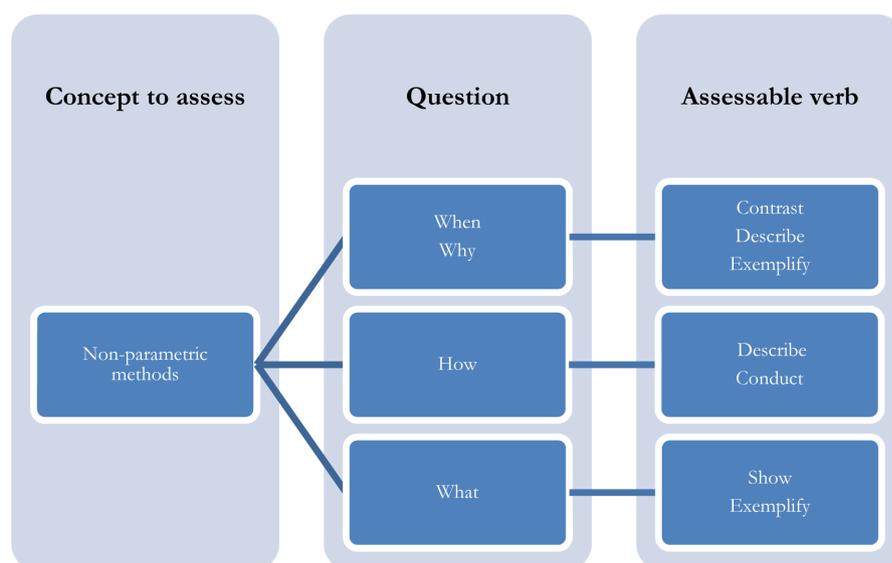


Figure 3. Design of learning outcomes in statistics.

relatively high transparency for the students, they cover the content from multiple perspectives (thus reducing the risk of engaging in surface learning only) and also ensure a high degree of assess ability.

3. Summary

This paper discusses various aspects of formulating relevant and assessable learning outcomes for undergraduate courses in statistics. It argues that clearly defined intended learning outcomes, specified within a course syllabus, outline or similar document, can provide a strong support for both the teacher and the students. Existing pedagogical models for learning, teaching and assessment of knowledge can provide a good support in this view but are at the same time too general to be used directly for. Instead, a narrower model is suggested, based on merging strong and deep verbs with detailed learning outcomes. The model is characterized by that the course literature has central role and that the matter of deep learning is involved already in the formulation of the learning outcomes. Specifically, it is proposed that the intended learning outcomes that the students are expected to attain after completing a course should (i) be clearly connected to the course literature to ensure their *relevance* and (ii) be associated with appropriate active *verbs* which in turn may be selected with the support of a two-dimensional graph that helps identifying verbs that are strong (=assessable) and deep (=stimulates deep learning). An explicit example is included that shows that appropriately specified learning outcomes should be easy to translate to exam questions. The suggested procedure should be helpful for teachers and course coordinators who develop new courses and also provide relatively high transparency for the students in the sense of being able to grasp the content of the course and what is being expected of them in terms of the final examination.

References

- American Association of Law Libraries. www.aallnet.org
- Anderson, L. W., Krathwohl, D. R. et al. (Eds.) (2001). *A Taxonomy for Learning, Teaching and Assessing*. Longman.
- Baume, D. (2002). *First Words*. Oxford Brookes University: OCSLD.
<https://www.brookes.ac.uk/services/ocslld/resources/assessment.html>
- Biggs, J. (1987) *Student Approaches to Learning and Studying*. Hawthorn, Vic: Australian Council for Educational Research.
- Biggs, J. (1993). What Do Inventories of Students' Learning Process Really Measure? A Theoretical Review and Clarification. *British Journal of Educational Psychology*, 83, 3-19. <http://dx.doi.org/10.1111/j.2044-8279.1993.tb01038.x>
- Biggs, J. B., & Collins, K. F. (1982). *Evaluating the Quality of Learning. The SOLO Taxonomy*. London: Academic Press.
- Biggs, J. B., & Tang, C. (2007). *Teaching for Quality Learning at University*. Maidenhead: McGraw Hill.
- Bland, M. (2000). *An Introduction to Medical Statistics* (3rd ed.). Oxford: Oxford University Press.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook I. The Cognitive Domain*. New York: David McKay Co Inc.
- Cox, D. R., & Hinkley, D. V. (1974). *Theoretical Statistics*. London: Chapman & Hall.
<http://dx.doi.org/10.1007/978-1-4899-2887-0>
- Crawley, E. F. (2007). *Rethinking Engineering Education: The CDIO Approach*. New York: Springer.
- Elmgren, M., & Henriksson, A. S. (2010). *Universitetspedagogik*. Stockholm: Nordstedts.
- Howitt, D., & Cramer, D. (2014). *Introduction to Statistics in Psychology* (5th ed.). Harlow: Pearson Education Limited.
- Marton, F., & Säljö, R. (1976a) On Qualitative Differences in Learning—1: Outcome and Process. *British Journal of Educational Psychology*, 46, 4-11. <http://dx.doi.org/10.1111/j.2044-8279.1976.tb02980.x>
- Marton, F., & Säljö, R. (1976b). On Qualitative Differences in Learning—2: Outcome as a Function of the Learner's Conception of the Task. *British Journal of Educational Psychology*, 46, 115-127.
<http://dx.doi.org/10.1111/j.2044-8279.1976.tb02304.x>

Appendix

The procedure is summarized in a flow-chart below:

Step I. List the content of the course by selecting headings from the Content index of the course literature.

Step II. Couple a number of action verbs (say 3 - 5) to each subject in the content list.

Step III. Place the action verbs, for each subject individually, in a two-dimensional Deepness/Strongness graph. Discharge verbs that are vague + not deep, and assure that at least some verbs are strong + deep.