ASSESSMENT PRACTICES AND STUDENTS' APPROACHES TO LEARNING: A SYSTEMATIC REVIEW

Jihan Rabah, eConcordia, Concordia University, Canada
&
Robert Cassidy, Concordia University, Canada
&
Manasvini Narayana, Concordia University, Canada

ABSTRACT

Student Approaches to Learning (SAL) differentiates between student learning objectives/ behaviors that are focused on the memorization of course content (surface approach) or the construction of meaning and comprehension (deep approach). For nearly five decades, researchers have explored SAL as an important framework for understanding, evaluating and guiding learning and instruction. An evidently robust and generally relevant construct, SAL has spawned numerous instruments to capture its Deep and Surface approaches, most notably the Study Process Questionnaire (SPQ; Biggs, 1976) and its successor the Revised Two-Factor SPQ (R-SPQ-2F; Biggs, Leung, Kember, 2001). Despite its history, there are few comprehensive reviews of its utility. Here we present the results of a systematic review of the literature describing how deep and surface approaches to learning are associated with different assessment practices.

This paper uses vote counting to investigate the relationship between assessment practices and students' approaches to learning. After a systematic search of the literature over 1,482 abstracts were reviewed, from which 21 articles were selected, and 53 voting scores were extracted. Several assessment types were grouped and analyzed to help explain the voting results. Pedagogical implications and suggestions for further research are discussed.

INTRODUCTION

SAL is one of the most enduring and robust constructs of the education research literature. Its surface and deep approaches differentiate between learning objectives and behaviors that focus on the memorization of course content or the construction of meaning and comprehension, respectively. Assessment practices, a powerful component of course design, influence learning expectations and the approaches students adopt in a course. Summative assessments are typically high stakes; post hoc assessments of learning and formative assessments offer scope for feedback and learning within the course. Both types of assessments could nudge students towards deep or surface approaches, based on their design. A systematic review of four decades (1976 – 2017) of research literature on SAL was performed, and the peer-reviewed articles that examined the association between assessment practices and student approaches were compiled and examined carefully for design quality and results. We present the results of this review describing how different student approaches to learning are associated with different assessment practices. The role of assessment in setting expectations for learning will be discussed.

RESEARCH QUESTIONS

The systematic review is an exploratory study that aims at answering the following research questions:

RQ1: What is the relationship between Deep Approach to Learning (as measured by the SPQ or revised version of the SPQ) and assessment practices?

RQ2: What is the relationship between Surface Approach to Learning (as measured by the SPQ or revised version of the SPQ) and assessment practices?

THEORETICAL FRAMEWORK

We are anchoring the systematic review in students' approaches to learning as measured by Biggs' Study Process Questionnaire (SPQ) and/or subsequent versions of it.

DATA SOURCES, EVIDENCE, OBJECTS, OR MATERIALS

When we started exploring the relationship of SAL as measured by SPQ or subsequent versions of it, we wanted our data sources to be comprehensive. In light of that, we explored various peer-reviewed databases. These included ERIC database, Canadian Business & Current Affairs, Academic Search Complete, PsycINFO, Proquest Dissertation, EdITLib, Communication & Mass Media Complete, and Medline. Our search strategies also included unpublished materials such as theses and research reports to avoid publication bias. We conducted web searches using several search engines such as Google and Bing to complement the data search process utilizing branching techniques to find as many articles as possible. All searches included a combination of SPQ key terms that varied according to researched database or source. These included the following terms "two-factor study process questionnaire," "two-factor study process questionnaire," "study process questionnaire," SPQ, RSPQ, R-SPQ, "R SPQ," R-SPQ-2F. The search was aimed at the abstract field of each database. Different databases required search key terms or search locations to be changed slightly to fit the appropriate Boolean search. Adaptations were made to the above keywords as needed.

Overall, 1,482 abstracts were reviewed. This initial number went down to 259 by the end of the first filtering phase[1]. The count went down further to 228 after removing duplicates and applying the following inclusion-exclusion criteria:

- The population being tested had to be university learners.
- Adoption of the SPQ or any subsequent version of it as a measure of student approach to learning.
- English version of the tool
- English as the language of publication of the article

Next, we categorized the 228 articles according to Biggs' three-P model: articles that related SAL to Presage, Process or Product of learning. We then decided to narrow in on studies that involved Presage characteristics. Presage was chosen specifically since it has repercussions on Product and Process of learning and can typically be manipulated to possibly improve the quality of teaching-learning.

This finally left us with 166 articles. The Presage category itself allows for more granular categorization. While clearly each of the presage components affects the other, we considered it prudent to categorize our articles in the hope of a more nuanced understanding. For example, while course design and assessment are more a continuum, we chose to analyze them separately, starting first with assessment practices. In the end, we had 21 studies in the assessment category hence 21 studies were included in this systematic review, and our research questions were: What is the relationship between Deep Approach to Learning (as measured by the SPQ or revised version of the SPQ) and assessment practices? What is the relationship between Surface Approach to Learning (as measured by the SPQ or revised version of the SPQ) and assessment practices?

METHODS AND MODES OF INQUIRY

Like with any systematic review, we followed a predefined sequence of steps in order to ensure transparency and replicability, following Cooper's (2016) approaches to systematic review reporting. The steps consist of the following:

1. Determine and develop the terms of reference of the research question.
2. Establish criteria for inclusion and exclusion of studies.
3. Develop a search strategy for identification of relevant studies.
4. Select studies based on abstract review.

[1] *This sharp drop can be explained by the fact that SPQ is used as an acronym for a couple of other instruments in business, psychology, and psychiatry fields.*

5. Select studies based on full-text review.
6. Code study features and moderator variables.
7. Conduct analysis and interpretation.

Inter-rater reliability was established to assure the validity of the extracted information: two coders worked independently and rated 10% of each article batches at each point- screening- eligibility- inclusion and coding for study features. At all times percent agreement was 88% or higher. When disagreements between coders occurred, mismatches in rating were discussed and a final agreement reached, to tighten procedures of acceptance and rejections. Each study was combed for instances of associations between the variables identified in the two research questions. Individual voting results and relationships between variables of interest were recorded onto Excel spreadsheets.

VOTE COUNTING PROCEDURES

Our data set included 21 studies dated between 1976 and 2016. Research countries included Australia, Greece, Netherlands, Belgium, and Canada. Research methodologies varied and incorporated highly statistical studies as well as descriptive and mixed methods studies. For a comprehensive observation of relations between assessment practices and learning approaches, qualitative as well as quantitative results were extracted from the different studies. To maximize the includes of the qualitative as well as quantitative studies, we had to be agile in choosing our methodology, and hence vote-counting methodology was used. A voting mechanism for each type of finding was charted. Instead of effect sizes, a categorization system was created to code the strength of each association. For coding purposes, these associations were categorized according to data type. Labeled A to F, they included the following: test of difference, correlation, regression, proportion of sample, point on a Likert scale, and/or qualitative statement of the relationship. As an example, consider a study exploring the research question: what is the relationship between the form of assessment and learning approaches? If the reported r value was significantly positive, it received a score of +2. If it was positive but not significant, it got a score of +1.0 and was marked with no relationship, -1 for a negative but not significant relation, and -2 for a significant negative relationship.

RESULTS

For this analysis, and after we proceeded to answer the aforementioned research questions, findings were as follows:

RQ1: Relationship between Assessment practices and Deep Approaches to Learning. Our data set gave 26 votes counted, indicating an overall average vote of +0.58

RQ2: Relationship between Assessment practices and Surface Approaches to Learning. The data set gave 27 votes counted, indicating an overall average vote of -0.19

WEIGHING VOTING RESULTS BY QUALITY AND SAMPLE SIZE

We noticed that the quality of studies varied, and because it is recommended in systematic reviews to appraise the quality of studies (Huff, 2009; Petticrew & Roberts, 2006; Ramey & Rao, 2011), we coded for the quality of studies. Studies were coded for 'low' quality, 'medium' quality, and/or 'high' quality.

However, we realized that our studies varied substantially by sample size as well since we included qualitative as well as quantitative results. In light of that we recalculated weighted voting for quality and for sample size so that we get a more gauging figure of the voting results the qualitative and quantitative differences.

With regards to quality, the highest quality was coded 0.8. The latter was considered 1 and all the other quality values were divided by it to make sure they were related to 'the highest scoring quality value'.

With regards to weighted sample size, we considered the median of the sample size and worked with the assumption that the median sample size is a full score and anything above the media is also is a full score. For studies with sample sizes smaller than the median (N=107), we used a logarithmic method to calculate the impact of the sample size on the vote size. That means we took into effect the decreasing effect of the increase in sample size from 0 to the median value, which is 107. The resultant weighted voting for quality and sample size was calculated by multiplying the raw voting score by the resultant quality factor and the resultant sample factor. The final results were as follows:

RQ1: Relationship between Assessment practices and Deep Approaches to Learning. Our data set gave 26 votes counted, indicating an overall average vote of +0.44

RQ2: Relationship between Assessment practices and Surface Approaches to Learning. The data set gave 27 votes counted, indicating an overall average vote of -0.28

THEME-BASED ANALYSES BY ASSESSMENT TYPES

To add detail and access additional insights to the results of the vote counting, we took a qualitative approach to analyzing the themes. For a thematic analysis, we first clustered the articles based on the type of assessment involved in the study. The purpose was to see patterns in reported discussions if any, between SAL and specific assessment types. The 21 studies in our pool clustered as follows: Multiple Choice: 6, Portfolio: 4, Essay (Long form writing): 2

All the others studied individual approaches that did not cluster into any themes or groups (For example: there were single studies dealing with Multimedia assessments, Progress Testing, Viva, Case Study and so on).

In this process, we first listed all the key claims and observations made by the authors in the results, discussion and conclusion sections of the papers. The authors were seeking to explain better, add nuance, or justify their findings. Having listed these author observations, we then extracted themes from them. Some of the recurring themes are reported below, organized by assessment type.

MULTIPLE CHOICE QUESTIONS (MCQS) AS ASSESSMENT

It appears that students are flexible in the learning approaches they employ to succeed in MCQs. In some cases, students moved from a surface approach at one point to a deep approach at a later point, as in the case of Dickie's (1994) freshman physics students. In Leung's et al. (2008) study, however, students began with a deep approach and later moved to a surface approach. From student interviews, Leung et al. (2008) conclude this shift happens on account of a big workload - or more accurately because the students perceive their workload to be big. But in spite of this shift to surface approach, these students still did better than those who started out and stayed with surface approaches.

Yonker's (2011) study of psychology students found that students' surface approach harms their MCQ performance more than deep approach helps them. This is contrary to common practitioner belief that MCQs encourage and reward a surface approach. The impact of student perception played a role again, in this case: when students

perceived MCQs as something that tested lower cognitive ability, they tended to take a surface approach, which in turn lead them to a poorer performance on the MCQs.

Rajaratnam et al. (2013) find that students with deep approach did very well on an MCQ Exam, in the physiology context of their study. The authors also use the findings to justify the use of didactic teacher-driven instructional methods, since the students were exposed to such a method before they attempted the MCQ exam.

To summarize the three themes in the MCQ studies: Relying purely on surface approach does not help students perform well on MCQ exams. Taking a deep approach might help, or not, but is less likely to harm performance. The workload is seen a factor driving students to surface approaches.

PORTFOLIO ASSIGNMENTS AS ASSESSMENTS

Studies that investigated learning approaches with Portfolio as an assessment type found that students' deep approach increases with higher order thinking tasks, and surface approach reduces. Gijbels et al. (2006) found that students with higher deep approach prefer portfolio assessments especially if they allowed the students to demonstrate their more complex learning. However, after exposure to several portfolio formative assessments, students tended to prefer the portfolio lesser. The authors speculate that the workload may have been a factor leading to this shift from deep to surface approach.

Segers et al. (2008) and Fong and Wai (2012) also find that portfolio increases the deep approach and decreases surface approach. Segers et al. (2008) add that however, approach to learning also relates to the student's perceptions of the portfolio task. The quality of feedback was seen as an important aspect in stimulating learning. Segers et al. (2008) identify several characteristics of feedback, which they believe is more amenable to nurturing deep approaches.

Baeten et al. (2008) found that even though students seemed to prefer deep approach when it came to portfolio tasks, it did not predict better performance. In his study, students shifted to surface approach on portfolio tasks, but that did not improve performance either. This author too speculates that the reasons for the shift are workload, and the students not being adequately motivated. He also emphasizes that students strategically shift between approaches based on their context. In summary, these studies seem to suggest that Portfolio assessments are quite likely to trigger deep approaches, but the deep approach in itself does not seem to be a predictor of success on portfolio tasks.

ESSAY /LONG FORM WRITING

In the case of essay type answers too, perception seems to matter. Gerzina at al (2003) found that students who perceived the essay answers as not being representative of the course content adopted a surface approach. If they did perceive the essays to be representative of the course, they adopted a deep approach.

As with MCQ, in the essay assessments, students seemed to use the two approaches strategically. Verkade and Lim (2016) found that the writing assessments did not require "deep writing" – a form of writing comparable to "deep reading." Even though students in his study took a deep approach to reading, they did not mimic the style of their reading in their writing. Verkade and Lim (2016) conclude that perhaps the test in question had an issue – it assumed that a deep approach for reading would translate into a deep approach for writing, though he does not get into too much detail about what "deep writing may specifically engender. In his study, the majority of the subject did not leverage comprehensiveness of reading, even though students were required to undertake comprehensive "deep" reading. Students in that study preferred a deep approach but were willing to apply a surface approach strategically.

From among MCQ, Portfolio and Essay assignments, if we were to distill themes further, it would be as follows: Strategic use of SAL, Student perception of assessment tasks affecting performance, the design of assessments affecting SAL and performance, and student workload (or perception of workload), and the role of feedback.

SCIENTIFIC OR SCHOLARLY SIGNIFICANCE OF THE STUDY OR WORK

While we recognize that this is a relatively exploratory systematic review, it does provide evidence that assessments do have a relation to students' approaches to learning. In conjunction with a review of studies involving course design and SAL, it might be possible to arrive at a deeper understanding of the relationships between Presage components and students' approaches to learning.

REFERENCE LIST

An asterisk marks studies included in the systematic review.

*Biggs, J. B. (1979). Individual and group differences in study processes. *British Journal of Educational Psychology*, *48*(3), 266-279.

Biggs, J. B. (1987). *Study Process Questionnaire manual. Student approaches to learning and studying.* Hawthorn, Australia: Australian Council for Educational Research Ltd.

Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71(1), 133-149.

*Baeten, M., Dochy, F., & Struyven, K. (2008). Students' approaches to learning and assessment preferences in a portfolio-based learning environment. *Instructional Science*, *36*(5), 359-374.

*Barndt, R. J. (2011). *Effect of Learning Contracts and Examination Problem Types on Students' Mastery of Managerial Accounting Principles*. Unpublished doctoral dissertation. Widener University, New York, USA.

*Chen, Y., Henning, M., Yielder, J., Jones, R., Wearn, A., & Weller, J. (2015). Progress testing in the medical curriculum: students' approaches to learning and perceived stress. *BMC Medical Education*, *15*(1), 147.

*Cobb, K. A., Brown, G., Jaarsma, D. A., & Hammond, R. A. (2013). The educational impact of assessment: a comparison of DOPS and MCQs. *Medical Teacher*, *35*(11), e1598-e1607.

Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach* (Volume 2). Sage publications.

*Dickie, L. (1994). *Approach to Learning and Assessment in Physics*. John Abbott College, CP 2000, Ste Anne de Bellevue, Quebec H9X 3L9, Canada.

*Fong, C., & Wai, R. (2012). *Student perceptions and learning approaches of using blogs in IT education for reflection and knowledge construction*. Unpublished doctoral dissertation. Durham University, Australia.

*Gerzina, T. M., Worthington, R., Byrne, S., & McMahon, C. (2003). Student use and perceptions of different learning aids in a problem-based learning (PBL) dentistry course. *Journal of Dental Education*, *67*(6), 641-653.

*Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference? *Educational Studies*, *32*(4), 399-409.

Haggis, T. (2003). Constructing images of ourselves? A critical investigation into 'approaches to learning' research in higher education. *British Educational Research Journal*, 29(1):89-104.

*Hamm, S., & Robertson, I. (2010). Preferences for deep-surface learning: A vocational education case study using a multimedia assessment activity. *Australasian Journal of Educational Technology*, *26*(7).

*Hillard, R. (1995). *The Effects of Examinations and the Assessment Process on the Learning Activities of Undergraduate Medical Students*. Unpublished doctoral dissertation. University of Toronto, Toronto, Canada.

Huff, A. S. (2009). *Designing for Research for Publication*. Thousand Oaks, CA: SAGE.

*Kiguwa, P., & Silva, A. (2007). Teaching and learning: addressing the gap through learning styles: short reports: teaching and learning in psychology. *South African Journal of Psychology*, *37*(2), 354-360.

* Leung, S.F., Mok, E., & Wong, D. (2008). The impact of assessment methods on the learning of nursing students. *Nurse Education Today*, *28*, 711-719.

*Mathiopoulos, K. & Paraskeva, F. (2009). Authentic Assessment in Computer Supported Collaborative Learning Environments in Higher Education. In T. Bastiaens, J. Dron & C. Xin (Eds.), *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2912-2916). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

*Nkhoma, M., Sriratanaviriyakul, N., Pham Cong, H., & Khai Lam, T. (2014). Examining the mediating role of learning engagement, learning process and learning experience on the learning outcomes through localized real case studies. *Education & Training*, *56*(4), 287-302.

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford, UK: Blackwell.

*Rajaratnam, N., Cruz, S. M., & Chandresakar, M. (2013). Correlation Between The Learning Approaches Of First Year Medical Students And Their Performance In Multiple Choice Questions In Physiology. *NJIRM*, *4*(5), 43-48.

Ramey, J., & Rao, P. G. (2011, October). The systematic literature review as a research genre. Professional Communication Conference (IPCC), *2011 IEEE International* (pp. 1–7, 17–19). Cincinnati, Ohio, USA.

*Smith, S. N., & Miller, R. J. (2005). Learning approaches: Examination type, discipline of study, and gender. *Educational Psychology*, *25*(1), 43-53.

*Segers, M., Gijbels, D., & Thurlings, M. (2008). The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies*, *34*(1), 35-44.

*Verkade, H., & Lim, S. H. (2016). Undergraduate Science Students' Attitudes Toward and Approaches to Scientific Reading and Writing. *Journal of College Science Teaching*, *45*(4), 83.

Watkins, D. (1983). Depth of processing and the quality of learning outcomes. *Instructional Science*, 12:49-58.

*Weller, J. M., Henning, M., Civil, N., Lavery, L., Boyd, M. J., & Jolly, B. (2013). Approaches to learning for the ANZCA Final Examination and validation of the revised Study Process Questionnaire in specialist medical training. *Anaesthesia and Intensive Care*, *41*(5), 631.

*Yonker, J. E. (2011). The relationship of deep and surface study approaches on factual and applied test-bank multiple-choice question performance. *Assessment & Evaluation in Higher Education*, *36*(6), 673-686.

Contact email: jihan.rabah@econcordia.com