

Flipped Classroom Evaluation using the Teaching Analysis Poll

Kristin Vogelsang¹, Frank Ollermann²

¹Osnabrueck University, Germany, ²University of Applied Sciences Osnabrueck, Germany.

Abstract

The increasing digitalization is a challenge for universities. Therefore, electronic teaching is experiencing a boost. Flipped Classroom is a special form of digital teaching is. Here, the basics are taught via online content and are interactively deepened in presence phases in class. For this quite modern variant of blended learning, there are numerous case descriptions in research with different evaluations available. However, a systematic examination of the evaluation of this form of teaching has not yet been carried out. The aim of this paper is, therefore, to diminish this research gap and to present the model of the Teaching Analysis Poll (TAP) as a systematic and easy to handle evaluation method in the Flipped Classroom. After a short review of recent research in the corresponding areas, we give a systematic overview of evaluations of Flipped Classroom. We introduce from the example of a Flipped Classroom Course the application of TAP as an adequate evaluation format.

Keywords: *Flipped Classroom; TAP; Evaluation; Case Study.*

1. Introduction

Digitalization is a major challenge of our time, which is why the use of digital media plays a special role in university teaching. The implications of digital teaching events are discussed in many ways in academia (Bender, 2016). In addition to advantages such as independence from time and space and ubiquitous availability, disadvantages such as the reduction of contact times, loss of control in the transfer of knowledge and the loss of important meta-competencies are also mentioned as arguments for and against digital teaching (Jacot, Noren, & Berge, 2014). A Flipped Classroom (FC) provides the advantages of digital teaching and avoids the disadvantages. Within a flipped or inverted classroom, the basic content is taught at home via online material (e.g., videos) and the in-class time is used for the training of deeper understanding and application. As the concept is rather new and still developing, a special emphasis should be put on the quality of education and permanent improvement. Systematic evaluations could achieve this goal (Scriven, 1996). The number of research articles dealing with the topic is massively increasing during the last years. Many of them focus on case description and analysis (Bishop & Verleger, 2013). Often information about the evaluation of the FC is provided. There is a great variety of FC approaches and at the same time no consistency regarding the way how to evaluate the FC classes (Vogelsang & Hoppe, 2018). This is why there is still a need for an evaluation procedure that enables a comparison and common understanding of the different FC approaches while at the same time doing justice to the heterogeneity of FC approaches. The paper aims at the systematic examination of the teaching analysis poll (TAP) as a method for systematic FC evaluation that seems to show great promise for fulfilling these requirements. A TAP is a structured, moderated group discussion, aiming at a qualitative evaluation of the quality of the support provided to the learning process. We will define the concepts of TAP and FC and give a summary of actual FC evaluations based on findings from the literature. Afterwards, we will introduce the TAP as a method for FC evaluation. We demonstrate the feasibility by an undergraduate class. Furthermore, we will present and discuss the results and mention conclusions as well as the limitations of the research.

2. Teaching Analysis Polls for Evaluation

The Teaching Analysis Poll has its roots in the tradition of midterm evaluations that developed in response to the often unsatisfactory end-of-semester evaluations that mainly provide quantitative data and don't perceivably lead to teaching improvements, especially from the current students' perspectives who often don't believe that the evaluation results are taken seriously by the teachers (Frank, Fröhlich, & Lahm, 2011). Clark and Redmond (1982) proposed Small Group Instructional Diagnosis (SGID) as a method to bring about formative feedback from the students to the teacher. TAP, as typically practiced in German universities, is a streamlined variant of SGID. The following is a usual sequence:

1. The teacher and the consultant (usually from the universities teaching center or the like) make an appointment for a course session during which the TAP is to be conducted.
2. The consultant conducts the TAP, usually during the last 30 minutes of the appointed session.
3. Each student answers three questions on a worksheet:
 - a. Which aspects of the course facilitate your learning?
 - b. Which aspects of the course impede your learning?
 - c. What can be done to overcome the learning obstacles?
4. The students gather in small groups to collect and discuss their answers.
5. The consultant collects the answers from all the groups and makes sure that everybody is clear about the meaning of each statement.
6. The consultant collects votes on the students' answers and notes the number of students who agree to the respective statement.
7. The consultant edits the results and sends a short written report to the teacher. This report usually contains only those statements that obtained a majority of votes during the classroom visit.
8. The teacher and the consultant meet for discussing the TAP results.

3. Flipped Classroom

Flipped Classroom (FC) is a teaching and learning paradigm that has its roots in several developments. One is the general tendency for teachers to focus more strongly on the students' actual learning processes than on the content and knowledge that is to be delivered by the lecturer. Barr and Tagg (1995) coined the corresponding slogan "Shift from Teaching to Learning" (even if they discussed this shift mainly from an organizational and economic perspective). A similar motto is King's (1993) "From Sage on the Stage to Guide on the Side", which prescribed a more intensive use of interactive deep-learning activities. Another root of the FC paradigm is the technical development, especially concerning web-based video technologies. This allowed teachers to use the video medium more and more for documenting and accompanying their classes and lectures by having them recorded and these recordings distributed over the internet, making it easier to reach their students as opposed to being restricted to video tapes or DVDs that have to be distributed physically and with much more effort. Basically, flipping or "inverting the classroom means that events that have traditionally taken inside the classroom now take place outside the classroom" (Lage, Platt, & Treglia, 2000, p. 32). Some authors go a little further and include the use of computer-based instruction as a defining aspect of FC (Bishop & Verleger, 2013, p. 5).

4. Flipped Classroom Evaluation in Practice

As the concept of Flipped Classroom is rather new, the success of FC lectures is still questionable. While most authors describe mainly positive impacts of the concept (Cui Tan, Wei-Gang Yue, & Yu Fu, 2017) negative estimations of the concept can also be found that assert no significant differences regarding learning outcomes when compared to traditional lectures (Gillette et al., 2018). To support their findings, researchers use different metrics to assess and evaluate the FC-concepts. But standards in FC evaluation are still rare (Stöhr & Adawi, 2018). A comparable overview is still missing and researchers. Researchers call for more standardized evaluation methods to assess the FC (O’Flaherty & Phillips, 2015). Our research contributes to closing the research gap about “How or why does the flipped classroom work, for whom, and in what circumstances?” (Stöhr & Adawi, 2018). This overview may help to cluster the great variety of FC assessment approaches and to lead to a better understanding of the purpose and implications.

To establish an evaluation for an FC concept, the *aim* must be defined. Formative evaluations monitor the students’ student’s learning to provide ongoing feedback for the instructors to improve their teaching and for students to improve their learning (Baker, 2014). To compare the FC to (mostly) traditional lecture designs (Hibbard, Sung, & Wells, 2016), often the summative assessment is used. Some studies combine both, measuring, e.g. the learning outcomes and perceptions about the lectures (Baker, 2014; Vogelsang & Hoppe, 2018). Most evaluations focus at the students (Lundin, Rensfeldt, Hillman, Lantz-Andersson, & Peterson, 2018). Evaluations measuring the lecturers’ perceptions and advance are rare (Baker, 2014).

Evaluations often compare performance measured in mid- and endterms (Amresh, Carberry, & Femiani, 2013). A pre-term assessment is useful if later comparisons or estimations about the development are desired (Al-Zahrani, 2015). Furthermore, accompanying in-class assessments enable previous knowledge to be checked and quality to be assured (Mahoney, Zappe, Butler, & Velegol, 2015). Midterm assessments are a common technique to evaluate learning success (Demski, 2013). In-class and online-assessments, complement the flip model (Mahoney et al., 2015). Some authors combine both methods (Sun, Xie, & Anderman, 2018). The decision about *when* to evaluate defines the possibilities of analysis. For an adequate evaluation design, it is also necessary to be sure *how* the evaluation should be conducted (Mahoney et al., 2015). Qualitative data is often used to assess self-reported perceptions about the class design while quantitative measures are often used for acceptance (Vogelsang & Hoppe, 2018) or efficacy (Sun et al., 2018) evaluations. They are often built on existing scales that measure perceptions, subjective experiences and satisfaction (Foldnes, 2016). There is a wide bandwidth in evaluations of *what* is assessed. Attitudes towards the FC like enjoyment, self-confidence and perceived value of the content (Foldnes, 2017) as well as the learning success (McLean, Attardi, Faden, & Goldszmidt, 2016) and the effectiveness (Sun et al., 2018) can be evaluated. These last results mainly aim

at the exam grades. Further positive statements regarding the problem-solving abilities of students exist (Gillette et al., 2018). Despite the measurable learning success, Foldnes (2017) proves that the increase of group interaction positively influences the learning outcomes. Due to the variety of possible and existing FC evaluations, there is a clear lack of applicable and comparable standardized methods to evaluate FC (Vogelsang & Hoppe, 2018). FC is rich in both ways: knowledge mediation and interaction. Therefore a “rich” and qualitative approach can be useful for evaluation purposes. Many evaluation tools are complicated, time-consuming and are often based on premises that do not necessarily apply to FC. Following the call for standard evaluation method, we will in the following present the TAP as an evaluation method for FC scenarios.

5. TAP for a Flipped Classroom Scenario: A Case Study

The FC learning concept is applied in a learning group of about 40 students of the study programs “Information Systems” and “Business Administration”. The subject of the course is “business process management”. The concept consists of three different pillars: online-tutorials, face-to-face reading seminars and guest lectures from practitioners. Fundamental learning contents are made available to the students via video tutorials. The video tutorials replace the teacher-centered lecture format (Herreid & Schiller, 2013). Sets of slides which contain fundamental knowledge content are verbally described and narratively complemented by carefully chosen examples and applications. For the exchange of material, the digital learning management system Stud.IP is used. The second component of the FC environment is a so-called “reading seminar”. The students prepare different texts at home and prepare them for the in-class time where the texts are discussed. A professor moderates the discussion. Finally, guest lectures by professionals working in the industry and consulting form the third component of the course concept. This additional information completes the knowledge of the field by illustrative application examples in addition to the fundamentally structured video tutorials and the in-depth reading seminar.

6. Results TAP

The TAP was conducted at the end of the term in the absence of the teacher and according to the proceedings described in section 2. A total of 17 students participated in the TAP and generated six positive statements, five negative statements and one suggestion for improvement (Table 2 summarizes the results). The most important positive statements relate to aspects not specific to FC. Still, a solid majority of student value the flexibility offered by the online videos provided within the FC approach. The most striking negative statement, as well as the corresponding suggestion for improvement, are about a central aspect of FC,

namely the way the online videos are incorporated into the overall instructional design of the course.

Table 1: Results of the TAP.

a) Positive statements		N	%
Availability of trials and old exams		16	94 %
Lecture presentation files »slides«		15	88 %
Online videos (repetitive viewing possible)		14	82 %
Free time management		13	76 %
Discussion of the papers		8	47 %
PossibilityAbility to ask questionsend questions via e-mailemail		4	24 %
b) Negative statements		N	%
Little incentive to watch videos regularly and promptly (risk of »bulimia learning«)		14	82 %
Contents of the papers are not always relevant to the exam		10	59 %
Video volume is very low		7	41 %
Rare interaction of the course professor		6	35 %
Video style »read out« fast; the videos provide no »lecture feeling«)		3	18 %
c) Suggestion for improvement		N	%
Meetings in between the videos (e. g. to work on assignments, concrete tasks in the sense of a tutorial)		17	100%

7. Discussion and Outlook

The results of the TAP underline the fit of TAP for FC courses. TAP, more systematically than standard evaluation questionnaires, aims at formative evaluation (*aim*), as TAP reflects the form of the lecture. By using standardized questions, both target groups, students and lecturers, are prompted to reflect their behavior and improve in the future (*who*). In our case, we conducted the TAP at the end of the term (*when*). For future evaluations, we would strongly favor a midterm evaluation. This allows the lecturer to react to the results and to adapt the teaching behavior, which can be a positive signal regarding the needs of students and may lead to an improved and more trustful teaching situation. FC, as well as TAP, allows teachers to learn about their students, their perspectives, learning processes, etc. We regard this mutual influence as a major asset of this evaluation, as students are forced to reflect upon

their learning behavior, outcomes, and perceptions (*what*), which is much less pronounced in the case of standardized summative evaluation questionnaires. FC as well as TAP focus on the learning process. This is why this kind of evaluation fits the situation very well. FC prepares students for active engagement in discussion and reflection, as prescribed by TAP. The students feel responsible for their learning process as the FC demands more self-regulation. As TAP is more a qualitative evaluation (*how*), there is no standard scale for interpretation available (*scale*). Using TAP consequently may nevertheless lead to a pool of results that allow for a systematic classification and interpretation. Future research may contribute to a common collection of TAP results aiming at a common pool of results and the measures taken afterwards. Possible research may also examine combinations of standard evaluations and TAP for FC, allowing for a comparison between these two evaluation approaches. Furthermore, we would encourage researchers to compare the TAP findings of FC classes with those of a standard lecture control group. Despite the merits of our research, the process suffers from some limitations. Due to the shortness of the research paper, we were only able to present the findings using one case. This is why the example cannot be free of influences due to the sample and special situation. The expansion of our research in the future will counteract this limitation.

References

- Al-Zahrani, A. M. (2015). From passive to active: The impact of the flipped classroom through social learning platforms on higher education students' creative thinking. *British Journal of Educational Technology*, 46(6), 1133–1148.
- Amresh, A., Carberry, A. R., & Femiani, J. (2013). Evaluating the effectiveness of flipped classrooms for teaching CS1. In 2013 IEEE Frontiers in Education Conference (FIE) (pp. 733–735). <https://doi.org/10.1109/FIE.2013.6684923>
- Baker, E. (2014). Flipped Classroom learning outcomes in intro to is class: Looking at student outcomes and professor performance evaluations. *Proceedings of the AIS SIG-ED IAIM 2014 Conference*.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning—A new paradigm for undergraduate education. *Change: The Magazine of Higher Learning*, 27(6), 12–26.
- Bender, C. (2016). Die Vorlesung. Ein Auslaufmodell? *Forschung Und Lehre*, Heft 8, 692–694.
- Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *ASEE National Conference Proceedings*, Atlanta, GA (Vol. 30, pp. 1–18).
- Clark, D. J., & Redmond, M. V. (1982). Small group instructional diagnosis.
- Cui Tan, Wei-Gang Yue, & Yu Fu. (2017). Effectiveness of flipped classrooms in nursing education: Systematic review and meta-analysis. *Chinese Nursing Research*, (4), 192–200.
- Demski, J. (2013). 6 Expert tips for flipping the classroom. *Campus Technology*, 26(5), 32–37.

- Foldnes, N. (2016). The flipped classroom and cooperative learning: Evidence from a randomised experiment. *Active Learning in Higher Education*, 17(1), 39–49. <https://doi.org/10.1177/1469787415616726>
- Foldnes, N. (2017). The impact of class attendance on student learning in a flipped classroom. *Nordic Journal of Digital Literacy*, 12(01–02), 8–18.
- Frank, A., Fröhlich, M., & Lahm, S. (2011). Zwischenauswertung im Semester: Lehrveranstaltungen gemeinsam verändern. *Zeitschrift Für Hochschulentwicklung*.
- Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., & Broedel-Zaugg, K. (2018). A Systematic Review and Meta Analysis of Student Pharmacist Outcomes Comparing Flipped Classroom and Lecture. *American Journal of Pharmaceutical Education*. <https://doi.org/10.5688/ajpe6898>
- Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching*, 42(5), 62–66.
- Hibbard, L., Sung, S., & Wells, B. (2016). Examining the Effectiveness of a Semi -Self-Paced Flipped Learning Format in a College General Chemistry Sequence. *Journal of Chemical Education*, 93(1), 24–30. <https://doi.org/10.1021/acsjcherned.5b00592>
- Jacot, M. T., Noren, J., & Berge, Z. L. (2014). The Flipped Classroom in Training and Development: Fad or the Future? *Performance Improvement*, 53(9), 23–28. <https://doi.org/10.1002/pfi.21438>
- King, A. (1993). From sage on the stage to guide on the side. *College Teaching*, 41(1), 30–35.
- Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30–43.
- Lundin, M., Rensfeldt, A., Hillman, T., Lantz-Andersson, A., & Peterson, L. (2018). Higher education dominance and siloed knowledge: a systematic review of flipped classroom research. *International Journal of Educational Technology in Higher Education*, 15(1), 1. <https://doi.org/10.1186/s41239-018-0101-6>
- Mahoney, E., Zappe, E., Butler, S., & Velegol, S. (2015). The Evolution of a Flipped Classroom: Evidence-Based Recommendations. *Advances in Engineering Education*, 4(3), 1–37.
- McLean, S., Attardi, S. M., Faden, L., & Goldszmidt, M. (2016). Flipped Classrooms and Student Learning: Not Just Surface Gains. *Advances in Physiology Education*, 40(1), 47–55.
- O’Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education*, 25, 85–95. <https://doi.org/10.1016/j.iheduc.2015.02.002>
- Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151–161.
- Stöhr, C., & Adawi, T. (2018). Flipped Classroom Research: From “Black Box” to “White Box” Evaluation. *Education Sciences*, 8(1), 22. <https://doi.org/10.3390/educsci8010022>
- Sun, Z., Xie, K., & Anderman, L. H. (2018). The role of self-regulated learning in students’ success in flipped undergraduate math courses. *The Internet and Higher Education*, 36(Supplement C), 41–53. <https://doi.org/10.1016/j.iheduc.2017.09.003>

Vogelsang, K., & Hoppe, U. (2018). Development of an Evaluation for Flipped Classroom Courses. Proceeding of Multikonferenz Der Wirtschaftsinformatik (Mkwi), 821–832.