



HAL
open science

Technology-Enhanced Formative Assessment: Does Changing the Way Students Confront Their Views Change Their Behaviour and Outcomes?

Franck Silvestre, Jean-François Parmentier

► To cite this version:

Franck Silvestre, Jean-François Parmentier. Technology-Enhanced Formative Assessment: Does Changing the Way Students Confront Their Views Change Their Behaviour and Outcomes?. 2021. hal-04179720

HAL Id: hal-04179720

<https://hal.science/hal-04179720>

Preprint submitted on 10 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Technology-Enhanced Formative Assessment: Does Changing the Way Students Confront Their Views Change Their Behaviour and Outcomes?

Franck Silvestre¹ ^a and Jean-François Parmentier² ^b

¹*Université Toulouse 1 Capitole, IRIT, Toulouse, France*

²*IPSA Toulouse, IRIT, Toulouse, France*

franck.silvestre@irit.fr; jean-francois.parmentier@irit.fr

Keywords: technology-enhanced formative assessment, formative assessment processes, peer evaluation, processes comparison, Tsaap-Notes, Elaastic


Abstract: Tsaap-Notes and Elaastic are two web platforms designed to help teachers orchestrate formative assessment sequences based on a two-votes approach. The second vote is carried out by students after they have had the opportunity to compare their points of view. Elaastic has been designed with new requirements to improve the Tsaap-Notes process: to shorten the number of phases in the process and to collect information on misconceptions provided by students. These requirements have changed the process significantly: the students' activities during the phase of confrontation of points of view of each process are substantially different. In this paper we present the results of the comparison of the two processes on the basis of the same questions that were asked by the same teachers either with Tsaap-Notes or with Elaastic in the same courses but in different school years. The results indicate that there are no significant differences between the two processes regarding the transitions between the first and second vote and the outcomes obtained by the students at the second vote. These results encourage us to continue our research on Technology-Enhanced Formative Assessment using the data collected by the Elaastic platform.


1 INTRODUCTION

Formative assessment has been identified as a key factor to improve learning (Black and Wiliam, 1998; Sadler, 1998; Bennett, 2011). The theory of formative assessment described by Black & William highlights strategies and activities that characterize formative assessment sequences (Black and Wiliam, 2009). Several technological platforms provide support for the orchestration of formative assessment sequences during face to face or distance courses. For example, Audience Response Systems, allows the implementation of different processes from most basic ones (simple question-answers loop) to more complex ones, like the Peer Instruction introduced by Mazur (Crouch and Mazur, 2001), leveraging more or less strategies and activities identified by Black and William. Today, web platforms such as Kahoot! (Dellos, 2015), Poll Everywhere (Sikarwar, 2015) and many others find their place in classrooms where students can use their own device to participate what Beatty and Gerace

(2009) call Technology Enhanced Formative Assessment (TEFA) activities. Spector et al. (2016) recommend to "support the development and deployment of powerful formative assessment tools that align with 21st century skills especially in the areas of critical thinking and inquiry- and problem-based learning".

Tsaap-Notes and Elaastic are both web platforms supporting complex formative assessment processes. The process implemented in Tsaap-Notes (Pr1), described in more details in section 2, has been designed to (1) reflect strategies and activities described by Black and Wiliam (2009) and (2) to collect the best answers given by students in order to generate revision tests with textual feedbacks (Silvestre et al., 2015b). The process implemented in Elaastic (Pr2) is a trial to improve the one implemented in Tsaap-Notes. Pr2 was designed to match new requirements: reducing the time of sequences driven by the tsaap-notes process and allowing the teacher to collect more information about misconceptions expressed by students. To reach this goal, strong modifications on Pr1 had been made including changes in students' activities and core algorithms. We were then interested in the eventual side effects of such big changes

^a  <https://orcid.org/0000-0002-1134-8200>

^b  <https://orcid.org/0000-0000-0000-0000>

in the process. This lead us to the following research questions :Do changes in Pr2 introduce differences in terms of students' behaviour and outcomes during the activities orchestrated with elaastic ?

The paper is structured as follows: in section 2 we present in depth each phase of both tsaap-notes and elaastic processes and then highlight the differences between the two processes. In section 3 we present the method we used to address our research questions. Section 4 describes the analyse we carried out on our datasets and then, in section 5 we present and discuss the obtained results. We end with conclusion and perspectives.

2 PROCESSES DESCRIPTION

2.1 Tsaap-Notes workflow

The process implemented in Tsaap-notes has been designed to promote, in a context of mass education, the strategies and activities characterizing a sequence of formative evaluation identified by Black and Wiliam (2009):

- sharing success criteria with learners;
- questioning the class;
- evaluation in the form of comments rather than numerical scores;
- self-evaluation and peer-evaluation;
- the use of summative tests for formative use.

The process implemented in Tsaap-Notes is composed of 4 phases. During the first phase, the students answer a question asked by their teacher. This question can be a closed-ended question or an open-ended question. In the case of a closed question, the students have an interface allowing them to provide a written rationale to argue their choice. Students are also asked to indicate a degree of confidence in their answer using a 5-levels Likert scale.

When the teacher feels that enough students have completed this first activity, the second phase is triggered. Each student is then invited to study a rationale from another student. The presented rationale comes from a student who expressed a high level of confidence in a different choice. This confrontation of points of view is anonymous: the student does not know the author of the contradictory argument. Each student then has an interface allowing him/her to submit a new answer. The student can then modify his or her initial choice and argumentation. Figure 1 is a screen capture of the second phase interface provided by Tsaap-Notes for students.

The third phase consists of a peer evaluation: each student is asked to give an assessment in the form of a Likert scale at 5 levels on 3 explanations corresponding to the correct answer. The aim of this phase is, on the one hand, to provide a peer assessment activity and, on the other hand, to identify the best arguments proposed by the students so that the best evaluated ones can be recycled in the form of textual feedback in online revision tests (Silvestre et al., 2015b). The last phase is devoted to the restitution of the results and is an opportunity for the teacher to lead a discussion with the students.

Tsaap-Notes has been used successfully in different STEM courses (MSc of computer science, Undergraduate Physics) to engage students and to put them in reflexive postures in way observed by kind of similar formative assessment processes like the Peer Instruction one. The generated tests based on questions asked by the teacher and completed by textual feedbacks coming from students were largely used: more students actually played the generated tests than the one participating in the original formative assessments orchestrated with Tsaap-Notes (Silvestre et al., 2015a).

In the rest of the paper, we are only interested in closed-ended questions that quantitatively measure the change of opinion between the first and second phase.

2.2 Elaastic workflow

Elaastic is a derived application from Tsaap-Notes in which the process has been altered to match some new requirements: reducing the time of sequences driven by the tsaap-notes process and allowing the teacher to get more information on rationales corresponding to a wrong choice. To reach these requirements, the second phase of the process has been redesigned: each student indicates with the help of a 4-levels Likert scale his or her degree of agreement on a set of contributions provided by the other students. The set of contributions can contain up to 5 different arguments and contains explanations for both the correct and incorrect answers, regardless of the student's initial answer. Presented rationales are no more selected function of the level of confidence of the author in order to guaranty that each rationale will have a significant number of grades. The number of contributions assessed by each student is specified by the teacher at the start of the sequence. Figure 2 is a screen capture of the second phase interface provided by Elaastic for students.

This approach (1) condenses the Elaastic's process into 3 phases against 4 in the tsaap-note process, and

Master/Informatique/DL/2013_2015/DCLL/1 13 janvier 2015 15:32

Here is an alternative response to yours.
Please examine this response and then feel free to change your response, explanation or confidence degree if necessary.

Oui
Le javascript étant exécuté côté client, l'utilisateur a une copie, la contamination du GPL force donc la possibilité d'accès au code à l'utilisateur.

Alternative answer

La licence GPL et Javascript

Soit L une librairie javascript distribuée sous licence GPL . Soit E une entreprise qui modifie la librairie L pour en faire une version L2. L'entreprise E propose à des utilisateurs son application serveur qui utilise la version L2 de la librairie. Est ce que les utilisateurs peuvent exiger de récupérer le code source de la version L2 du logiciel ?

Oui
 Non

Give an explanation to your choice

En tant qu'utilisateurs de l'application serveur, on se retrouve dans le même cas que précédemment, on n'a pas de copie du logiciel. On ne peut donc toujours pas exiger récupérer le code.

What is the degree of confidence in your response (1: not confident to 5:very confident) ? 3

Submit

Second submission form: a chance to modify the first answer

Reply Favorite Learn

Figure 1: Phase 2 with Tsaap-Notes

(2) provides the teacher with grades given by students on all rationales corresponding to both correct and incorrect answers.

Table 1 summarizes the different phases of the 2 processes.

3 METHOD

We worked with 2 datasets: one coming from the Tsaap-Notes application, the other coming from the Elaastic application. A set of questions played with Tsaap-Notes has been played with the Elaastic platform by the same teachers in the same program played at different academic years. We used these set of questions to compare the two processes. For each sequence played with the 2 platforms, we were interested in the following indicators:

- number of correct and incorrect answers in the first vote (respectively R1 and W1),

- number of correct and incorrect answers in the second vote (respectively R2 and W2)
- number of transitions between the 2 votes from incorrect to correct, correct to incorrect answers (respectively WR and RW),
- number of students with a correct answer to the first and second vote (RR)
- number of students with an incorrect answer to the first and second vote (WW)

We considered that the sequences played with both Tsaap-Notes and Elaastic could only be used to compare the 2 processes if they had been played on comparable groups of students. The students interviewed belonged to the same course but were not questioned in the same academic year. We considered the groups to be comparable when, for a question asked with Tsaap-Notes and Elaastic, the differences in the results obtained on the first vote were not significant. We used a Chi-squared test and retained only those sequences for which the test provided a p-value greater

Comparing viewpoints

Here are presented one or several alternative responses. Please indicate how much you agree with these answers.

Choice [1]
 x^2 returns the same result for any number and it's opposite. Therefore $g(x)$ and $g(-x)$ returns the same result and that rule applies when $x = 4$ aswell. So if $g(4) = 8$ then $g(-4) = 8$
 Your evaluation:

I'm not giving my opinion 1 2 3 4 5

Choice [1]
 By solving the equation, if
 $g(4) = 8$
 $\Leftrightarrow a * 4^2 + 24 = 8$
 $\Leftrightarrow a = -1$
 Which means that $g(-4) = -1 * (-4)^2 + 24 = 8$
 Your evaluation:

I'm not giving my opinion 1 2 3 4 5

Choice [2]
 If $g(4)$ returns 16 then $g(-4)$ returns -16.
 Your evaluation:

I'm not giving my opinion 1 2 3 4 5

Take a second chance to change your answer and confidence degree.

Your answer : 1 2 3 4

Submit

Second submission form: a chance to modify the vote

Figure 2: Phase 2 with Elaastic

than 0.4.

Table 2 shows an example of what we done for each question asked with either with Tsaap-Notes or Elaastic. In this example, the p-value of 0.7066 indicates there is no significative difference between the group of students having played the sequence with Tsaap-Notes and those having played the sequence with Elaastic.

We thus obtained 7 Tsaap-Notes sequences comparable to 6 Elaastic sequences. 317 students participated in one of these sequences. 183 students played Tsaap-Notes sequences against 134 who played Elaastic sequences.

4 RESULTS AND DISCUSSION

Table 3 shows all the indicator values collected for the sequences being compared. Table 4 shows same indicators expressed in percentage.

When performing a Chi-square test on the contingency table obtained by filtering table 3 to keep only R1 and W1 columns, we obtain the following result: X-squared = 0.33646, df = 1, p-value = 0.5619. This

result confirms there is no significant difference between groups coming from Tsaap-Notes sequences or Elaastic sequences. This is what we expected given the way we selected comparable sequences. Proportions of correct answers are 69.40% for Tsaap-Notes and 65.67% for Elaastic, with a mean difference of 0.04 (95% confidence interval is [-0.07, 0.14]).

When performing a Chi-square test on the contingency table obtained by filtering table 3 to keep only R2 and W2 columns, we obtained the following result: X-squared = 0.22762, df = 1, p-value = 0.6333. This result indicates that there is no significant difference between groups after they participated second phase either with Tsaap-Notes or Elaastic. Proportions of correct answers are 75.41% for Tsaap-Notes and 72.39% for Elaastic, with a mean difference of 0.03 (95% confidence interval is [-0.07, 0.13]).

To infer that both processes are not significantly different, we have to check if Tsaap-Notes and Elaastic students changed their vote (or not) between phase 1 and 2 in the same way. To look at this trait, we performed a Chi-square test on the contingency table obtained by filtering table 3 to keep only transition columns (RR, RW, WR, WW). We obtained the following result: X-squared = 1.7421, df = 3, p-value

Table 1: Comparison between phases in Tsaap-Notes and Elaastic

	Tsaap-Notes	Elaastic
Phase 1	Each student answers a question giving written rationale and confidence degree	
Phase 2	Each student consults exactly one rationale provided by a peer and corresponding to a different answer.	Each student evaluates up to five rationales provided by a peers and corresponding to correct or incorrect answers.
	Each student submits a new response with the ability to change the initial choice and the argumentation.	Each student submits a new response with the ability to change the initial choice.
Phase 3	Each student evaluates 3 argumentations corresponding only to the correct answer.	NR
Last Phase	Restitution of results and discussion	

Table 2: Example of contingency table used to select comparable sequences

Question: In Scrum, any change can impact the development team at any moment. True or False ?		
	R1	W1
Tsaap-Notes	20	9
Elaastic	10	7

Pearson's Chi-squared test with Yates' continuity correction:
 $X\text{-squared} = 0.14171$, $df = 1$, $p\text{-value} = 0.7066$

= 0.6276. The p-value indicates that we cannot consider that students behave significantly differently in the two processes.

All these results indicate that Tsaap-Notes and Elaastic processes lead students in similar behavior and lead to comparable score after phase 2 when groups are comparable regarding their performance in phase 1. We thus consider that what is proposed in Phase 2 inside each process can be considered as an equivalent treatment to improve scores at the second vote.

To sum up, first, providing students with a single counter-rationale provided with a high level of confidence against several rationales with higher behavioural engagement (peer review requested) does not significantly change transitions and outcomes at the second vote and, therefore, does not appear to significantly change students' cognitive engagement. Second, the Elaastic process, at equal gain, offers advantages in relation to its initial requirements: the number of phases is reduced and it provides teachers with information given by students about their misconceptions.

5 CONCLUSIONS AND PERSPECTIVES

Tsaap-Notes and Elaastic implement two formative evaluation processes based on the principle of 2 votes for a question. The second vote takes place after students have confronted their points of view by evaluating alternative rationales. The Elaastic process is derived from the Tsaap-Notes process: it allows students to assess more than one rationale from the first vote corresponding to both correct and incorrect votes. Thus, the Elaastic process offers the opportunity to provide students with formative activities and, at the same time, to collect data on their misconceptions, in particular by allowing the identification of those that are rated the highest by students.

The study presented in this paper shows that there are no significant differences in terms of student behaviour and outcomes between the two processes. Therefore, because of the advantages of the Elaastic process compared to the Tsaap-Notes process, we are encouraged to further study the data resulting from the use of Elaastic by teachers and students. We wish to focus our work on the study of misconceptions perceived as convincing by students. We hypothesise that Elaastic allows us to identify the main misconceptions of students on the different topics addressed by teachers through Elaastic activities.

The application allows teachers to consult all the rationales provided by the students with the mean grade each of them receives during the peer evaluation. How teachers use or do not use this information is also a future direction for our work. Our aim will be to make it easier to take this information into account in the feedback phase or in a broader remediation phase at the instructional design level.

Table 3: Indicators gathered for both platforms based on comparable sequences

	R1	W1	RR	RW	WR	WW	R2	W2	Students count
Tsaap-Notes	127	56	115	12	23	33	138	45	183
Elaastic	88	46	82	6	15	31	97	37	134

Table 4: Indicators in percentage gathered for both platforms based on comparable sequences

	R1	W1	RR	RW	WR	WW	R2	W2	Student count
Tsaap-Notes	69.40	30.60	62.84	6.56	12.57	18.03	75.41	24.59	100
Elaastic	65.67	34.33	61.19	4.48	11.19	23.13	72.39	27.61	100

REFERENCES

- Beatty, I. D. and Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education and Technology*, 18(2):146–162.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in education: principles, policy & practice*, 18(1):5–25.
- Black, P. and Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1):7–74.
- Black, P. and Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1):5.
- Crouch, C. H. and Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American journal of physics*, 69(9):970–977.
- Dellos, R. (2015). Kahoot! a digital game resource for learning. *International Journal of Instructional technology and distance learning*, 12(4):49–52.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in education: principles, policy & practice*, 5(1):77–84.
- Sikarwar, A. S. (2015). Flipped classroom with poll everywhere: engaging students with active learning in large group settings. *Journal of Asian Scientific Research*, 5(2):111.
- Silvestre, F., Vidal, P., and Broisin, J. (2015a). Online tests based on contributions provided by teachers and students during face to face lectures. In *2015 IEEE 15th International Conference on Advanced Learning Technologies*, pages 29–33. IEEE.
- Silvestre, F., Vidal, P., and Broisin, J. (2015b). Reflexive learning, socio-cognitive conflict and peer-assessment to improve the quality of feedbacks in online tests. In *Design for Teaching and Learning in a Networked World*, pages 339–351. Springer.
- Spector, J. M., Ifenthaler, D., Sampson, D., Yang, J. L., Mukama, E., Warusavitarana, A., Dona, K. L., Eichhorn, K., Fluck, A., Huang, R., et al. (2016). Technology enhanced formative assessment for 21st century learning.