



**HAL**  
open science

# La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants

Jean-François Parmentier, Franck Silvestre

## ► To cite this version:

Jean-François Parmentier, Franck Silvestre. La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants. 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain - EIAH 2019, ATIEF : Association des Technologies de l'Information pour l' Education et la Formation, Jun 2019, Paris, France. hal-02904155

**HAL Id: hal-02904155**

**<https://hal.science/hal-02904155>**

Submitted on 21 Jul 2020

**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.

# La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants

Jean-François Parmentier<sup>1</sup> and Franck Silvestre<sup>2</sup>

<sup>1</sup> Université de Toulouse, UPS, IRES, F-31400 Toulouse  
jf.parmentier@gmail.com

<sup>2</sup> Université Toulouse 1 Capitole,  
Institut de Recherche en Informatique de Toulouse (IRIT)  
118 Route de Narbonne, F-31062 Toulouse CEDEX  
franck.silvestre@irit.fr

**Résumé.** Le processus en N Phases implanté dans la plateforme Tsaap-Notes permet une mise en oeuvre dans un contexte de cours en face à face d'une variante de l'instruction par les pairs. Or son utilisation dans un enseignement hybride, mêlant face à face et distance, a mis en évidence une limitation forte : un nombre significatif d'étudiants ne mènent pas le processus jusqu'à son terme. Cet article présente son adaptation pour une utilisation dans des enseignements hybrides ou à distances. L'adaptation principale consistant à ne plus synchroniser la transition entre les phases est inspirée de la plateforme DALITE. Les résultats de l'expérience menée pour mettre à l'épreuve ce nouveau processus démontrent clairement l'augmentation du nombre d'étudiants réalisant les N phases de bout en bout.

**Mots-clé:** Instruction par les pairs · expérimentation · pédagogie active · évaluation formative · évaluation par les pairs · enseignement à distance · enseignement hybride · analyse des données d'apprentissage · Tsaap-Notes · DALITE

**Abstract.** The N Phases process implemented in the Tsaap-Notes platform provides an implementation in a face-to-face course context of a variant of peer instruction. However, its use in hybrid teaching, combining face-to-face and distance, has highlighted a strong limitation : a significant number of students do not complete the process. This article presents its adaptation for use in hybrid or distance learning. The main adaptation consisting in no longer synchronizing the transition between phases is inspired by the DALITE platform. The results of the experiment conducted to test this new process clearly demonstrate the increase in the number of students performing the N phases from start to finish.

**Keywords:** Peer Instruction · experiment · active learning · formative assessment · peer assessment · distance learning · blended learning · learning analytics · Tsaap-Notes · DALITE

## 1 Introduction

Parmi les différentes pratiques de pédagogies actives, le processus d'évaluation formative baptisé « Instruction par les Pairs » (IP) concerne les enseignements dispensés en face à face dans des amphithéâtres pouvant accueillir plusieurs centaines d'étudiants. L'IP a été introduit par Mazur pour la première fois en 1991 durant des cours de physique pour des étudiants de licence [8]. L'approche consiste, pour une même question fermée (principalement des questions à choix exclusif), à sonder les étudiants deux fois : une première fois après une phase de réflexion individuelle, puis une deuxième fois après que les étudiants aient débattu avec leurs voisins sur le bien-fondé de la réponse qu'ils ont fournie à l'issue du premier sondage. Les résultats du deuxième sondage révèlent alors un plus grand nombre de bonnes réponses. Différentes expérimentations ont démontré les bénéfices de cette approche en termes d'apprentissage par rapport à une approche classique purement transmissive [8,6,14]. Devant les résultats probants apportés par l'IP tant sur l'engagement des étudiants que sur leurs résultats d'apprentissage, beaucoup d'universités se sont lancées dans l'utilisation des systèmes de votes individuels (tels des boitiers), en appliquant ou non l'IP [7,10,9].

La plateforme Web Tsaap-Notes [12] permet l'orchestration de séquences d'évaluation formative structurées de manière similaire à celles de l'IP. L'application, conçue pour une utilisation durant les cours en face à face, introduit l'argumentation écrite par les étudiants, la confrontation de points de vue anonymisée et l'évaluation par les pairs. Le processus supporté par Tsaap-Notes, a été baptisé processus en N Phases [13] ; les expérimentations menées en présentiel dans différents cours ont montré des résultats semblables à ceux obtenus en instruction par les pairs traditionnelle : fort taux d'engagement des étudiants et améliorations des bonnes réponses à l'issue du deuxième vote [12].

Parmi les enseignants utilisateurs de Tsaap-Notes, certains ont souhaité exploiter l'outil en dehors de leur enseignement en présentiel. Les premiers retours informels convergeaient vers le constat suivant : un nombre significatif d'étudiants ne prenaient pas part à la deuxième phase du processus, celle d'évaluation par les pairs. En conséquence, le nombre d'évaluations des arguments écrits était faible sur la plupart des contributions.

Ces retours nous ont conduits à nous interroger sur la pertinence de l'utilisation du processus en N Phases à distance et à poser la question de recherche suivante : sous quelle(s) condition(s) le processus N Phases conserve-t-il son caractère engageant dans un contexte d'activité menée en dehors du temps de classe ?

L'étude de l'application DALITE, conçue pour la mise en œuvre à distance du processus d'IP, nous a conduits à émettre deux hypothèses :

1. l'implantation d'un processus à transitions non synchronisées, tel celui existant dans DALITE permettant aux étudiants de participer aux différentes phases en une seule connexion à l'application, réduira de manière significative le taux de perte d'étudiants entre la première et la deuxième phase ;

2. l'augmentation du nombre d'étudiants à la deuxième phase permettra, quasi mécaniquement, d'augmenter le nombre d'évaluations de chaque contribution.

Afin de valider ces hypothèses, nous avons adapté le processus en N Phase et modifié Tsaap-Notes en conséquence puis mis en place un dispositif expérimental permettant de comparer les résultats obtenus entre l'utilisation de la version initiale du processus et l'utilisation de la version adaptée.

L'article est composé de cinq parties. Une revue bibliographique des logiciels permettant d'orchestrer les processus d'instruction par les pairs et dérivés est présentée dans la section 2. Le processus en N Phases adapté pour une utilisation hors classe est présenté dans la section 3. L'expérimentation et les résultats obtenus pour mettre à l'épreuve nos hypothèses sont détaillés respectivement dans les sections 4 et 5. Enfin, la section 6 conclut l'article en présentant les perspectives ouvertes par les résultats obtenus.

## 2 Logiciels pour l'orchestration des processus IP et dérivés

Une revue bibliographique a permis d'identifier l'existence de deux logiciels permettant l'orchestration des processus d'IP et dérivés : Tsaap-Notes et DA-LITE. Les deux plateformes sont présentées de manière détaillée avant d'être comparées.

### 2.1 Tsaap-Notes

La plateforme Tsaap-Notes [11] a été conçue pour une utilisation pédagogique en face à face. Son objectif initial est de proposer une amélioration de l'instruction par les pairs sur deux points :

- L'alignement pédagogique : dans sa forme initiale, l'instruction par les pairs utilise le débat oral. Les étudiants ne sont alors pas invités à formaliser par écrit leur raisonnement. Or, durant les évaluations sommatives, il est très fréquemment demandé aux étudiants d'expliquer par écrit les résultats qu'ils obtiennent. L'approche IP n'est donc pas optimale en termes d'alignement constructiviste [3] : l'activité proposée durant le cours pour atteindre l'objectif d'apprentissage n'est pas de même nature que celle proposée pour évaluer de manière sommative le niveau d'atteinte de l'objectif d'apprentissage.
- La connaissance des raisonnements des étudiants : lors de l'instruction par les pairs à l'aide de dispositifs de vote simples, l'enseignant ne dispose comme information que de l'histogramme des réponses. Or ce qui est attendu des étudiants n'est pas seulement de voter pour la bonne réponse, mais aussi d'être capable de justifier ce choix. Ainsi connaître les raisonnements des étudiants qui les ont conduits aux différents choix de réponses est crucial afin de pouvoir les corriger. En particulier, il est essentiel de détecter les faux positifs, c'est-à-dire les raisonnements faux

qui conduisent au bon choix de réponse, pour que les étudiants concernés ne repartent pas avec une fausse information sur la validation de leurs acquis. En général, dans la mise en œuvre de l'IP, l'enseignant tente d'obtenir cette connaissance en se déplaçant aléatoirement dans la classe en écoutant les débats entre élèves, ou en demandant à quelques étudiants d'expliquer publiquement leurs raisonnements. Or ces pratiques ont à la fois délicates à mener et peu représentatives.

Le processus en N Phases, dans sa forme générique, se déroule en 3 phases successives déclenchées par l'enseignant. Les transitions entre les phases sont donc synchronisées pour tous les étudiants. Les trois phases se déclinent de la manière suivante :

1. Chaque étudiant prend connaissance individuellement de la question et des choix proposés, puis effectue un choix et écrit une argumentation.
2. On propose alors à chaque étudiant d'évaluer jusqu'à trois explications proposées par ses pairs. Cette évaluation s'effectue en donnant une note de 1 à 5 (sous forme d'étoiles). Ils peuvent ensuite changer leur choix de réponse s'ils ont été convaincus par une des explications alternatives proposées (Fig. 1) [13].
3. La dernière phase propose à la fois une synthèse des votes sous forme d'histogramme, ainsi qu'un affichage des justifications les mieux notées. L'enseignant prend alors connaissance de quelques explications parmi les mieux notées et effectue des commentaires sur leur pertinence.

**Choix [5] :**  
Le plus important est de savoir ce qui les bloque pour pouvoir leur donner des explications pertinentes !  
★★★★★

**Choix [3] :**  
Il faut connaître ce que l'on enseigne, sinon on ne peut pas l'enseigner.  
★★★☆☆

Vous pouvez modifier votre sélection :

1    2    3    4    5

Figure 1. Phase 2 de TsaaP-Notes.

You answered **A** and gave this rationale :

The closer the ship, the sooner it gets hit !

Consider the problem again, noting the rationales below that have been provided by other students. They may, or may not, cause you to reconsider your answer. Read them, and select your final answer.

**A**

- Battleship A must get hit first, since it is closer
- They bove have aboute the same maximum height, so since A is closer it will get hit first.
- I stick with my own rationales

**C**

- The x-velocity never changes and the y-acceleration is the same on both so they will rise and fall simultaneously.
- The parabola of shell A has a different curbature than that of shell B, but the same x-intercepts.

Figure 2. Phase 2 de DALITE.

Dans la pratique, il est fréquent d'observer un raisonnement incorrect parmi les explications les mieux notées qui correspondent cependant au choix de la réponse correcte. Ainsi le procédé permet de détecter des faux positifs. De plus, l'enseignant peut consulter les explications les mieux notées correspondant à des choix de réponses incorrectes. La phase d'évaluation par les pairs donne

accès à l'enseignant à la connaissance des raisonnements « séduisants », donc représentatifs des conceptions erronées, sans que celui-ci n'ait besoin de lire l'ensemble des contributions écrites.

Les expérimentations menées dans différents cours en face à face ont montré des résultats semblables à ceux obtenus avec l'IP : fort taux d'engagement des étudiants et améliorations des bonnes réponses à l'issue du deuxième vote [12].

Ce processus a été étendu à l'usage de questions ouvertes. Dans ce cas, l'étudiant n'a pas de choix à sélectionner et ne fait qu'indiquer son raisonnement. La dernière phase n'affiche pas d'histogrammes. La validation s'effectue par les commentaires oraux de l'enseignant sur les différentes réponses qui ont été les mieux évaluées.

## 2.2 DALITE

DALITE est une application Web conçue pour exécuter à distance le processus d'Instruction par les Pairs [5]. Elle est donc destinée à un usage hors classe, soit en complément d'un enseignement présentiel [4,5], soit dans le cadre d'une formation 100% en ligne comme un MOOC [2].

La phase de débat, initialement réalisée à l'oral, a été adaptée à l'écrit. La plateforme a donc été conçue pour inciter les étudiants à s'expliquer par écrit à travers des justifications de QCM, à comparer ces justifications à celles de leurs pairs et à réfléchir à la qualité des justifications proposées. Dans sa version la plus récente [2], répondre à une question pour un étudiant se déroule en 3 phases successives :

1. Visualisation de la question, des choix, et écriture d'une argumentation.
2. Pour deux des choix de réponses, on propose à l'étudiant 3 ou 4 justifications écrites par d'autres étudiants. Ils doivent alors lire les explications puis sélectionner le choix et l'explication qu'ils pensent être les meilleurs. Ils peuvent décider de conserver leur propre explication (Fig. 2). L'un des deux choix proposés correspond toujours à la réponse correcte [4]. L'autre choix correspond soit à la réponse qu'a choisie l'étudiant (dans le cas où celui-ci a répondu incorrectement), soit à une réponse sélectionnée en amont par l'enseignant et correspondant habituellement à la réponse incorrecte la plus fréquente.
3. La dernière phase leur résume ce qui vient de se passer : leur choix et explication précédents, leur nouveau choix et leur nouvelle explication ainsi que l'explication donnée par un expert, mais sans leur indiquer à quel choix cela correspond [4,5,1].

De manière alternative, la phase 2 peut aussi se dérouler de manière séquentielle en proposant des choix de réponses un par un, et en demandant à l'étudiant de voter « pouce levé » ou « pouce baissé » pour chacun des choix [2].

Ainsi, les élèves participent à toutes les phases du processus dès leur première connexion à l'application. Les transitions entre les phases ne sont donc pas synchronisées d'un étudiant à l'autre. L'enseignant peut alors leur demander de compléter les activités en dehors des cours (en complément d'un enseignement

présentiel), ou au cours d'une formation en ligne, dans le cadre d'un MOOC par exemple.

Lors de la création d'une question, l'enseignant doit fournir [1] :

- la question sous forme QCM en indiquant la bonne réponse,
- une explication justifiant ce choix,
- le « deuxième meilleur choix » qui sera proposé aux étudiants ayant voté pour la bonne réponse.

L'outil a été utilisé dans différents cours de mécanique [5,1] pour des étudiants âgés de 17 à 19 ans. Les analyses ont montré qu'utiliser DALITE comme outil supplémentaire à un cours traditionnel conduit à une amélioration des gains conceptuels mesurés par un test standardisé de mécanique [4,5]. Les gains sont comparables à l'utilisation de l'instruction par les pairs en classe. Ainsi, DALITE semble pouvoir être utilisé comme une alternative au Peer Instruction en classe.

### 2.3 Comparaison des processus

Tsaap-Notes et DALITE proposent des fonctionnalités globalement similaires : réponse à des questions QCM sous forme d'écrit (phase 1 identique), confrontation de point de vue sous forme de visualisation des choix et argumentations des autres participants, changement de vote possible, et affichage final des histogrammes. Il y a cependant des différences significatives :

- **Le caractère synchronisé des transitions entre les phases :** DALITE fonctionne en transitions de phases non synchronisées : lorsqu'un étudiant se connecte, il effectue immédiatement les différentes phases indépendamment de ce que font les autres étudiants. Au contraire, Tsaap-Notes fonctionne en transition de phases synchronisées : les différentes phases sont déclenchées l'une après l'autre par l'enseignant.
- **Sélection vs Évaluations des contributions :** dans DALITE l'étudiant sélectionne 1 explication (qui peut-être la sienne) qui correspond à son second vote. Dans Tsaap-Notes l'étudiant doit évaluer chacune des contributions qu'on lui présente et peut changer ou non son vote en conséquence.
- **Informations pour l'enseignant :** dans DALITE, seuls les histogrammes des votes sont affichés à l'enseignant. Tsaap-Notes propose de plus l'affichage des explications les mieux notées.

Les différences constatées s'expliquent assez naturellement en comparant les objectifs de conception initiaux. L'utilisation de Tsaap-Notes en cours en face à face nécessite une maîtrise du temps passé sur chaque phase et donc une gestion synchronisée des cadences des étudiants. L'objectif d'échanges entre enseignant et étudiants sur les argumentations écrites explique une présentation des résultats incorporant les contributions des étudiants avec leur note moyenne issue de l'évaluation par les pairs, là où DALITE s'intéresse d'abord au changement d'avis entre les deux votes.

La non-synchronisation des transitions proposée par DALITE a directement inspiré la conception du processus en N Phases adapté à une utilisation à distance. Les principales caractéristiques du nouveau processus sont présentées dans la section suivante.

### 3 Processus en N Phases à transitions non synchronisées

#### 3.1 Principes généraux

Le nouveau processus permet à l'étudiant de réaliser l'ensemble de ses tâches lors de sa première visite sur l'application. L'étudiant est donc invité à répondre à la question posée en fournissant sa contribution textuelle ; après soumission de sa réponse, le système lui propose immédiatement d'évaluer un lot de contributions. L'étudiant peut modifier sa réponse si la confrontation de point de vue induite l'y pousse. À l'issue de cette étape, le processus se termine de deux manières différentes selon que l'activité se déroule dans un contexte hybride ou entièrement à distance.

Dans le contexte d'un enseignement délivré entièrement à distance, l'étudiant accède immédiatement au feedback proposé par la plateforme : son score à la question, la réponse et l'explication attendues, la note moyenne obtenue par sa contribution, les argumentations des étudiants les mieux notées ayant répondu correctement. Dans le contexte d'une utilisation hybride, l'étudiant accèdera au feedback sur son travail après que l'enseignant ait décidé de publier les résultats sur la question. Cette publication est en général effectuée pendant un cours en présentiel.

#### 3.2 Garantir l'accroissement du nombre d'évaluateurs

Le processus présenté dans la section 2.1 dispose de l'ensemble des contributions des étudiants pour effectuer l'affectation entre les contributions et les étudiants évaluateurs. Dans ce cas de figure, il est possible de générer le plan d'affectation à partir d'une analyse de l'ensemble des réponses. Un des premiers algorithmes utilisés par Tsaap-Notes pour réaliser le plan d'affectation est décrit de manière détaillée dans [13]. Dans le cas de notre nouveau processus, les affectations doivent être réalisées au fil de l'eau à partir des contributions disponibles à l'instant où l'étudiant accède à la plateforme. L'algorithme utilisé consiste à proposer à l'étudiant courant un lot d'au plus 5 contributions alternatives correspondant aux contributions qui ont été les moins évaluées par les pairs à cet instant. Cette approche garantit qu'à l'issue de l'activité, chaque contribution a été évaluée par un nombre maximum d'évaluateurs possible.

Cependant, l'algorithme pose un problème de démarrage à froid, car il ne fonctionne qu'en présence de contributions fournies par les pairs. Or, le ou les premiers étudiants connectés à la plateforme ne disposent pas de telles contributions. La solution imaginée pour régler ce problème est le suivant : l'enseignant, quand il prépare sa question, dispose d'une interface lui permettant de fournir la réponse attendue à la question ainsi qu'autant de réponses factices imaginées pour cette question qu'il le souhaite. Ces réponses fournies par l'enseignant sont alors utilisées par l'algorithme d'affectation pour le ou les premiers étudiants se connectant à la plateforme tant que des contributions issues des pairs ne sont pas encore disponibles.

## 4 Dispositif expérimental

Afin de mettre à l'épreuve les deux hypothèses formulées dans l'introduction, une expérimentation a été menée sur deux ans. La première année proposait à des étudiants d'utiliser Tsaap-Notes dans sa version initiale. La deuxième année, une version incluant le processus à transitions non synchronisées a été testée sur un autre groupe d'étudiants aux caractéristiques similaires.

La comparaison des deux processus a été effectuée en analysant les données d'apprentissage recueillies au regard de deux indicateurs :

1. Combien d'étudiants participent à la phase d'évaluation des contributions? L'objectif étant de maximiser la confrontation de points de vue afin d'améliorer l'apprentissage des étudiants.
2. Par combien d'étudiants est évaluée chacune des explications? L'objectif étant de garantir la pertinence des contributions qui sont sélectionnées pour permettre à l'enseignant de remédier aux conceptions erronées.

### 4.1 Population et organisation

Les deux expérimentations ont eu lieu au cours d'un enseignement d'initiation à l'algorithmique à destination des étudiants de L1 à l'Université Paul Sabatier, en 2016-2017 (groupe 1) puis en 2017-2018 (groupe 2), aux premiers semestres universitaires.

Les thématiques abordées et l'organisation du cours étaient similaires. Chaque semaine était composée d'un cours en amphithéâtre (CM) en classe entière d'une durée de 1h ainsi que de deux heures de travaux pratiques sur machines en petits groupes (quart de classe). Le nombre d'étudiants présents aux cours magistraux était similaire entre les deux années (en moyennes 73 et 69).

L'année 2016-2017, le cours en amphithéâtre avait lieu chaque vendredi et les étudiants devaient ensuite effectuer un Devoir Maison (DM) utilisant Tsaap-Notes. Avant le lundi soir, ils devaient répondre à trois questions ouvertes (phase 1 de contribution), puis ils devaient, entre le mardi et le prochain cours, évaluer les contributions de trois autres étudiants pour chacune des questions (phase 2 d'évaluation). Enfin, au début de chaque cours suivant, l'enseignant faisait des commentaires oraux sur certaines réponses des étudiants (phase 3 de présentation des résultats). Ce cycle a eu lieu 3 fois, conduisant donc à 3 DM.

Les mêmes questions ont été posées durant l'année 2017-2018. Les cadences étant cette fois-ci non synchronisées, ils avaient toute la semaine pour compléter leurs réponses et évaluer immédiatement après 3 autres contributions pour chacune des questions. Au début du cours suivant, de manière similaire à l'année précédente, l'enseignant effectuait des commentaires oraux sur quelques contributions et quelques évaluations.

Les questions posées portaient sur des aspects métacognitifs (réflexion sur leur propre apprentissage). Une question type et une réponse typique d'étudiant sont présentées dans le tableau 1.

**Table 1.** Exemple de question et de réponse d'étudiant.

<b>Question</b>	Quel était le point le plus important pour vous dans le cours de vendredi dernier (thème du 18 novembre) ? Expliquez pourquoi. Soyez précis et détaillé.
<b>Réponse typique</b>	"Le plus important est la notion d'optimisation des programmes pour qu'ils renvoient un résultat en moins de temps possible et avec un cout en mémoire plus faible. En cas de boucle, chercher à faire le moins de tours possible."

Deux types de données ont été analysées : le nombre d'étudiants complétant chacune des phases (écriture puis évaluation) pour chacune des questions de chacun des DM, puis le nombre moyen d'évaluations reçues par une contribution.

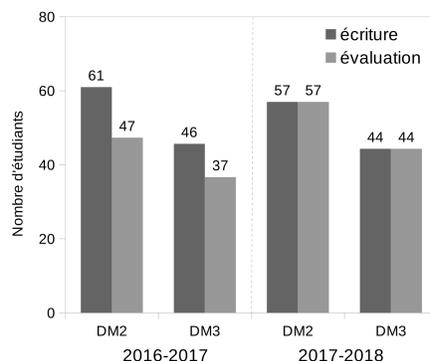
## 5 Résultats

Un nombre important d'étudiants hors de la population cible ayant effectué le DM1 de l'année 2016-2017, l'analyse des données s'est effectuée uniquement sur les DM 2 et 3 pour chacune des années. Au total 6 questions ont été posées. En moyenne 53 étudiants ont répondu à chacune des questions pour l'année 2016 et 51 pour l'année 2017, conduisant à un total de 624 contributions écrites.

La figure 3 présente les résultats pour chacun des DM. Les valeurs indiquées représentent la moyenne des contributions sur les 3 questions du DM correspondant. Par exemple, pour le DM2 de l'année 2016-2017, 60 étudiants ont répondu à la question 1, 59 à la question 2 et 64 à la question 3, conduisant à une moyenne de 61. Les valeurs pour chacune des questions étant relativement similaires au sein d'un même DM, le choix a été fait d'effectuer des moyennes par DM. Au contraire, on constate une diminution du nombre de participants entre le DM2 et le DM3, pour chacune des années. Celle-ci est de 25% pour l'année 2016 (intervalle de confiance à 95% : [14%, 37%]) et de 23% pour l'année 2017 ([13%, 36%]), l'écart entre les deux résultats étant non statistiquement significatif ( $p$ -value = 1 avec un test exact de Fisher). La proportion de participants aux DM par rapport au nombre moyen d'étudiants en cours est aussi similaire entre les deux années (83% pour le DM2 et 64% pour le DM3). Ces deux résultats confirment l'hypothèse d'une population similaire entre les deux années.

Une différence majeure entre les deux années est cependant observée dans le nombre de participants à la phase d'évaluation des contributions. Ainsi, on observe une perte moyenne de 21% (intervalle de confiance à 95% : [14%, 30%]) entre les deux phases durant l'année 2016 et une perte nulle ([0%, 4%]) durant l'année 2017. Cette différence est statistiquement très significative ( $p < 8 \cdot 10^{-8}$ ) (calcul sur le tableau 2). Les deux populations étant supposées identiques, cette différence de comportement est attribuée au résultat de la modification du processus.

Durant l'année 2016-2017, des interviews ponctuelles et informelles semblent montrer une acceptation et un intérêt de la part des étudiants à évaluer le travail



**Figure 3.** Nombre d'étudiants ayant participé aux phases d'écriture et d'évaluation pour chacun des DM. Les valeurs sont les moyennes sur les trois questions du DM.

**Table 2.** Nombre moyen d'étudiants participants aux différentes phases. Les participants sont ajoutés entre les différents DM (ex :  $101 = 57 + 44$ ).

	écriture et évaluation	écriture seule
<b>2016-2017</b>	84	23
<b>2017-2018</b>	101	0

des autres étudiants. Cependant les étudiants déclarent avoir des difficultés à organiser leur travail au cours de la semaine, expliquant un oubli régulier de participer à la phase d'évaluation. Ces déclarations semblent confirmer que la différence de comportement entre les deux années s'explique majoritairement par le changement de processus.

La perte d'étudiants observée en 2016-2017 diminue l'impact pédagogique du dispositif : moins d'étudiants participent à la phase d'évaluation, diminuant ainsi la qualité de l'évaluation moyenne de chaque contribution. En théorie, comme chaque élève doit évaluer 3 contributions, chacune de celles-ci est évaluée 3 fois, la moyenne obtenue donnant un indicateur relativement correct de la qualité de la contribution. Or en raison de la perte d'étudiants entre les deux phases, le nombre d'évaluateurs est plus faible que le nombre de contributeurs, conduisant à une moyenne de 2.2 évaluations par contribution durant l'année 2016-2017. Celle-ci est à 2.9 pour l'année 2017-2018. La moyenne du nombre d'évaluations est donc en nette augmentation et converge vers la moyenne attendue de 3. Le résultat est inférieur à 3 en raison de la présence des fausses contributions créées par l'enseignant et du nombre plus faible d'évaluations des 3 dernières contributions étudiantes. Les résultats sont synthétisés dans le tableau 3.

**Table 3.** Synthèse des données et des résultats

<b>Année universitaire</b>	2016-2017	2017-2018
<b>Présence moyenne en cours d'amphi</b>	73	69
<b>Transitions</b>	synchronisées	dissociées
<b>Participation moyenne (écriture)</b>	53	51
<b>Participation moyenne (évaluation)</b>	42	51
<b>Perte moyenne</b>	21 %	0 %
<b>Nombre moyen d'évaluations par contribution</b>	2.3	2.9

## 6 Conclusion et perspective

Le processus d'instruction par les pairs et ses dérivés assistés par la technologie sont reconnus pour les bénéfices qu'ils apportent aux étudiants durant leur apprentissage. En particulier, le processus en N Phases implanté dans Tsaap-Notes, favorise l'alignement entre les activités proposées en cours et celles proposées lors des évaluations sommatives et permet à l'enseignant d'accéder aux raisonnements des étudiants. Cependant, le processus en N Phases, comme l'IP, a été conçu pour une mise en œuvre durant les cours en face à face. Lors des premières expérimentations portant sur l'utilisation de l'application Tsaap-Notes dans un contexte d'enseignement à distance, l'analyse des données d'apprentissage a révélé deux dysfonctionnements : une perte non négligeable de participants à l'activité entre la première et la deuxième phase du processus ainsi qu'un nombre faible d'évaluateurs par contributions à l'issue de la phase de confrontation de point de vue et d'évaluation par les pairs. Ces résultats nous ont incités à concevoir et à implanter dans Tsaap-Notes une adaptation du processus en N Phases intégrant la non-synchronisation des transitions entre les phases telle qu'elle est implantée dans la plateforme DALITE. L'analyse des données d'apprentissage à l'issue des expérimentations faisant usage de la nouvelle version du processus, a mis en évidence la disparition de la perte de participants entre la première et la deuxième phase et l'augmentation significative du nombre moyen d'évaluateurs par contribution.

Bien que le nouveau processus apporte une réponse satisfaisante pour l'utilisation de Tsaap-Notes dans un contexte d'enseignement à distance, de nouvelles problématiques émergent. En effet, afin d'augmenter le nombre d'évaluateurs pour chaque contribution, l'algorithme initial utilisé par Tsaap-Notes a été modifié pour maximiser le nombre d'évaluateurs par contribution. En contrepartie, aucun autre critère n'est utilisé pour effectuer l'association entre contributions et évaluateurs. À titre d'exemple, la version initiale de l'algorithme s'appuyait sur l'indice de confiance des réponses pour planifier les appariements. Par ailleurs, le processus ne règle pas le problème de la "coupure à chaud" : à ce jour, le processus ne permet pas d'évaluer la contribution du dernier étudiant participant à l'activité. Ces problèmes guident aujourd'hui notre réflexion sur la conception de nouveaux processus.

## Références

1. Bhatnagar, S., Desmarais, M., Whittaker, C., Lasry, N., Dugdale, M., Charles, E.S. : An analysis of peer-submitted and peer-reviewed answer rationales, in an asynchronous Peer Instruction based learning environment. In : International Educational Data Mining. Madrid, Spain (Jun 2015)
2. Bhatnagar, S., Lasry, N., Desmarais, M., Charles, E. : DALITE : Asynchronous Peer Instruction for MOOCs. In : European Conference on Technology Enhanced Learning. Lyon, France (Sep 2016)
3. Biggs, J. : Enhancing teaching through constructive alignment. *Higher education* **32**(3), 347–364 (1996)
4. Charles, E.S., Lasry, N., Lenton, K., Whittaker, C., Dugdale, M., Bhatnagar, S. : Les réseaux conceptuels collectifs en enseignement et en apprentissage : l’usage de TIC pour relier la science scolaire avec la réalité extérieure (2014)
5. Charles, E.S., Lasry, N., Whittaker, C., Dugdale, M., Lenton, K., Bhatnagar, S., Guillemette, J. : Beyond and within classroom walls : Designing principled pedagogical tools for students and faculty uptake. In : Computer Supported Collaborative Learning (2015)
6. Crouch, C.H., Mazur, E. : Peer instruction : Ten years of experience and results. *American journal of physics* **69**(9), 970–977 (2001)
7. Fagen, A.P., Crouch, C.H., Mazur, E. : Peer Instruction : Results from a Range of Classrooms. *The Physics Teacher* **40**(4), 206–209 (Apr 2002). <https://doi.org/10.1119/1.1474140>
8. Mazur, E. : Peer instruction (1997)
9. Parmentier, J.F., Lamine, B., Bonnafé, S. : Changer les conceptions en mécanique des étudiants en L1 à l’Université. In : M01 Mini symp. Formation et pédagogie. AFM, Association Française de Mécanique, Lyon (2015)
10. Rudolph, A.L., Lamine, B., Joyce, M., Vignolles, H., Consiglio, D. : Introduction of interactive learning into French university physics classrooms. *Physical Review Special Topics - Physics Education Research* **10**(1), 010103 (Jan 2014). <https://doi.org/10.1103/PhysRevSTPER.10.010103>
11. Silvestre, F. : Conception et mise en oeuvre d’un système d’évaluation formative pour les cours en face à face dans l’enseignement supérieur. Thèse de doctorat, Université Paul Sabatier, Toulouse, France (novembre 2015), (Soutenance le 25/11/2015)
12. Silvestre, F., Vidal, P., Broisin, J. : 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, pp. 339–351. Springer (2015)
13. Silvestre, F., Vidal, P., Broisin, J. : Un nouveau processus d’évaluation pour améliorer la qualité des feedbacks dans les tests en ligne. *Sciences et Technologies de l’Information et de la Communication pour l’Éducation et la Formation* **24**(spécial), pp–181 (2016)
14. Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., Su, T.T. : Why Peer Discussion Improves Student Performance on In-Class Concept Questions. *Science* **323**(5910), 122–124 (Jan 2009). <https://doi.org/10.1126/science.1165919>