

# Proposition d'Indicateurs d'Écriture Collaborative et leur Évaluation

Fahima Djelil<sup>1</sup>[0000-0001-8449-2062], Christian Hoffmann<sup>2</sup>[000-0002-0620-3621], Anis M. Haddouche<sup>1</sup>[0000-0002-5321-3988], Nadine Mandran<sup>2</sup>[000-0002-8660-3827], and Cédric d'Ham<sup>2</sup>[0000-0002-7313-7097]

<sup>1</sup> IMT Atlantique, Lab-STICC, UMR CNRS 6285, 29238 Brest, France

{fahima.djelil,mohamed-anis.haddouche}@imt-atlantique.fr

<sup>2</sup> Université Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

{christian.hoffmann, nadine.mandran, cedric.dham}@univ-grenoble-alpes.fr

**Résumé.** L'écriture collaborative est une activité courante en éducation, faisant l'objet d'un champ de recherche en pleine croissance dans le domaine de l'Apprentissage Collaboratif Assisté par Ordinateur. Dans cet article, nous proposons deux indicateurs permettant de mesurer la contribution des apprenants à la rédaction collaborative d'un texte, appelés "balance de contribution" et "co-écriture". Nous nous intéressons également à la relation existante entre les mesures de ces indicateurs avec des stratégies d'écriture collaborative dites "sommatives" et "intégratives". De plus, nous évaluons l'interprétation de ces indicateurs et ces stratégies par des enseignants, en utilisant des documents texte rédigés dans une plateforme numérique collaborative appelée *LabNbook*. Les résultats montrent que les enseignants ont une bonne interprétation de ces indicateurs et stratégies.

**Mots-clé :** Indicateurs · Écriture Collaborative · Apprentissage Collaboratif Assisté par Ordinateur · Analyse de l'Apprentissage · LabNbook.

**Abstract.** Collaborative Writing is a common activity in education, which is a growing research field in Computer-Supported Collaborative Learning (CSCL). In this paper, we propose two indicators that provide measures of student contributions to a text collaborative writing, namely "balance of contribution" and "co-writing". We are also interested in the existing link between these indicators and collaborative writing strategies called "summative" and "integrative". Moreover, we proceed to the evaluation of the interpretation of the indicators and the strategies by teachers, using text documents written in a collaborative digital platform called LabNbook. Results showed that teachers have a good interpretation of these indicators and strategies.

**Keywords:** Indicators , Collaborative Writing , Computer Supported Collaborative Learning (CSCL), Learning Analytics , LabNbook.

## 1 Introduction

L'écriture collaborative suscite un intérêt de recherche croissant dans le domaine de l'apprentissage collaboratif assisté par ordinateur [4,26]. Une grande partie des travaux existants vise à comprendre la dynamique de collaboration entre apprenants, en s'appuyant sur l'analyse de traces numériques, ainsi que les interactions entre pairs tels que les échanges de messages écrits ou les conversations orales [26]. En dépit de l'abondance des travaux dans ce domaine, il existe un besoin accru en termes de mesures et de métriques pouvant servir à mieux comprendre l'écriture collaborative [26].

Dans cet article, nous proposons deux nouveaux indicateurs permettant de mesurer l'écriture collaborative : la *balance de contribution* reflétant dans quelle mesure les contributions des apprenants à un texte co-écrit sont équilibrées et la *co-écriture*, reflétant l'intensité des interactions entre apprenants. Leurs calculs reposent sur une mesure de dispersion basée sur la variance [9]. Nous explorons également le lien existant entre ces indicateurs et des stratégies d'écriture collaborative issues de la littérature. Les définitions de ces indicateurs résultent d'un travail antérieur, que nous ne détaillons pas dans cet article [10]. Ce travail est ancré dans un contexte réel et s'appuie sur une plateforme numérique collaborative, appelée *LabNbook* [6]. Par ailleurs, nous évaluons l'interprétation des indicateurs et des stratégies par des enseignants qui ne sont pas des utilisateurs experts de *LabNbook*. Les questions de recherche que nous abordons sont les suivantes : QR1) Quels sont les indicateurs (métriques) permettant de mesurer la contribution des apprenants en écriture collaborative; QR2) Comment déduire des stratégies d'écriture collaborative à partir de ces indicateurs; QR3) Dans quelle mesure ces indicateurs et ces stratégies sont-ils interprétables par les enseignants ?

Cet article est organisé comme suit. La section 2 présente un état de l'art sur l'écriture collaborative. La section 3 présente les indicateurs proposés ainsi que les stratégies d'écriture collaborative. La section 4 décrit notre méthode de recherche. La section 5 donne le calcul des indicateurs proposés et leur relation avec les stratégies d'écriture collaborative. Les résultats de leur évaluation sont discutés en Section 6, les conclusions et les perspectives sont données en Section 7.

## 2 L'écriture Collaborative

L'écriture collaborative fait référence à la production d'un texte par deux ou plusieurs auteurs (co-auteurs) [21]. Elle est définie comme un processus impliquant des interactions substantielles entre des apprenants partageant la prise de décision et les responsabilités afin de produire un document unique [14,21,25]. Les taxonomies et modèles existants [18,14,21] ayant pour but de caractériser l'écriture collaborative, témoignent d'un intérêt de recherche soutenu pour ce domaine. Certains modèles offrent un cadre permettant d'examiner les comportements et les stratégies d'écriture collaborative grâce à l'analyse de traces numériques [17,22,16,1,5,12,25].

On peut citer un modèle très connu dans le domaine de langue étrangère [21], fondé sur deux concepts, *l'égalité*, reflétant le niveau de contribution et de contrôle de l'apprenant sur une tâche donnée, et *la mutualité*, reflétant le niveau d'engagement d'un

apprenant sur la contribution d'un autre apprenant. Ces deux fondements s'apparentent à d'autres concepts définis dans des travaux antérieurs. À titre d'exemple, [7] soutient que la collaboration implique la négociation et met l'accent sur le degré de symétrie des interactions entre pairs. Le terme symétrie est emprunté au domaine des systèmes homme-machine et désigne une balance équilibrée de contrôle, induite par la collaboration entre un système et son utilisateur [8]. L'ajustement mutuel est une deuxième caractéristique de la collaboration par la négociation [3]. Il s'agit de stratégies mises en œuvre pour parvenir à un accord dans l'interaction [8]. Cela se reflète par les différentes façons d'éditer un texte écrit par d'autres individus et les façons d'interagir avec ceux qui éditent son propre texte [11]. Cela représente également le degré d'engagement avec les idées et les textes écrits par les autres et la participation à la production d'un écrit commun [13].

Ces travaux nous amènent à penser que l'écriture collaborative peut être caractérisée par deux types d'indicateurs. Le premier mesurant l'égalité des contributions des apprenants, et le second, le degré de leurs interactions.

### 3 Indicateurs et Stratégies d'Écriture Collaborative

#### 3.1 La Balance de Contribution et la Co-écriture

Nous définissons la *balance de contribution* comme une métrique qui mesure combien les contributions d'apprenants à un texte co-écrit sont égales, équilibrées ou déséquilibrées. Elle s'aligne sur celle proposée par [16], appelée *équilibre de participation*, mesurée en considérant 1 moins la variance des parts des membres d'un groupe dans le travail collaboratif. Dans ce travail [16], le choix de la variance a été justifié par sa simplicité et sa facilité d'interprétation comparée à d'autres méthodes de calcul tel que le coefficient de Gini ou l'indice de Blau. L'indicateur *balance de contribution* est calculé en se basant sur la métrique de variance, mesurant une distribution entre les contributions moyennes des auteurs à la rédaction d'un document en terme de mots.

L'indicateur de *co-écriture* mesure la contribution de l'apprenant, de la même manière que pour la balance de contribution, mais au niveau de la phrase. En effet, nous considérons qu'une phrase est porteuse d'une idée, et l'évaluation de la collaboration dans la rédaction des phrases nous renseigne sur la manière dont les apprenants interagissent lors de l'écriture collaborative. Alors que la balance de contribution mesure la division du travail (nombre de mots qu'un apprenant apporte en contribution à un texte), la co-écriture va au-delà en mesurant la co-construction du texte par les apprenants (modification et ajout de mots dans les phrases écrites par d'autres dans le texte). Les valeurs des deux indicateurs se situent entre 0 et 1 (voir section 5).

#### 3.2 Les Stratégies Sommativ et Intégrative

Les indicateurs proposés peuvent nous renseigner sur les stratégies d'écriture collaborative mises en œuvre par un groupe d'apprenants lors de la production d'un texte. Nous nous intéressons particulièrement aux deux stratégies parmi cinq proposées par [17], à savoir :

*i) construction séquentielle sommative de textes, i.e.* un membre du groupe présente un document qui constitue une proposition initiale, partielle ou complète pour la résolution d'une tâche, et les autres participants ajoutent successivement leurs contributions au document initial, sans modifier ce qui a été écrit précédemment, donc en acceptant systématiquement ce qui est ajouté par les autres co-auteurs.

*ii) construction séquentielle intégrative de textes, i.e.* un membre du groupe présente un document qui constitue une proposition de tâches initiales, partielle ou complète, et les autres membres du groupe contribuent successivement à ce document initial, en proposant des modifications justifiées ou en discutant de leur accord ou non avec ce qui a été écrit précédemment.

Nous utilisons le terme *sommative* pour désigner une stratégie selon laquelle chaque apprenant ajoute du texte sans modifier le texte d'autres apprenants, et selon laquelle le résultat est une juxtaposition des contributions individuelles; et le terme *intégrative*, pour désigner une stratégie selon laquelle un apprenant propose une version initiale d'un texte et les autres apprenants contribuent successivement en apportant des modifications au texte existant [10]. Les apprenants n'adoptent pas nécessairement une stratégie unique mais une combinaison des deux.

## 4 Méthode de Recherche

Notre objectif consiste, d'une part, à construire des indicateurs du niveau de collaboration des apprenants, en s'appuyant sur l'analyse de traces numériques recueillies lors de la production d'un texte par les apprenants, et d'autre part, à vérifier que ces indicateurs sont compréhensibles par les enseignants. Pour ce faire, nous adoptons la méthode de posture épistémologique du Constructivisme Pragmatique (CP) [2], dont la construction des hypothèses de la connaissance scientifique est basée sur la prise en compte de la réalité et de l'humain dans cette réalité. Nous avons choisi la méthode CP car dans ce travail, nous avons besoin de données de terrain produites par les opinions des enseignants pour évaluer les métriques proposées. Nous nous appuyons également sur la méthode Recherche Orientée par la Conception (Design-Based Research) [24] ainsi que les guides associés [15] pour mener notre recherche. Cette méthode propose, entre autres, de construire les connaissances et les outils associés de manière itérative en intégrant les acteurs du terrain.

### 4.1 Contexte

En vue d'évaluer nos indicateurs, nous utilisons *LabNbook*, une plateforme numérique pour l'apprentissage des sciences expérimentales [6]. Elle fournit des outils pour la rédaction de documents scientifiques collaboratifs et permet aux apprenants d'interagir entre eux. De plus, *LabNbook* fonctionne en mode "co-édition verrouillée" [23], c'est-à-dire que les apprenants peuvent travailler simultanément dans l'espace de travail partagé mais chaque document (appelé *LabDoc*) qui le compose ne peut être édité que par un seul apprenant à la fois.

## 4.2 Collecte de Données

Afin d'évaluer dans quelle mesure ces stratégies et ces indicateurs sont interprétables par les enseignants, nous avons administré un questionnaire à 15 enseignants, utilisateurs non-experts de la plateforme *LabNbook*. Nous avons sélectionné 12 LabDocs, chacun est co-écrit par 2 à 4 élèves dans des conditions réelles (activités en classe). Les contributions à l'écriture des différents apprenants ont été mises en évidence. Nous avons choisi des LabDocs ayant différentes stratégies d'écriture : 4 LabDocs selon une stratégie intégrative, 4 LabDocs selon une stratégie sommative, et 4 LabDocs selon des stratégies mixtes. Les 15 enseignants ont examiné les 12 LabDocs, un à la suite de l'autre. Après avoir lu les définitions des stratégies et des indicateurs et visualisé chaque LabDoc, ils indiquent :

- (1) la stratégie d'écriture parmi 6 choix : EI (Entièrement Intégrative), PI (Plutôt Intégrative), ESI (Entre Sommative et Intégrative), PS (Plutôt Sommative), ES (Entièrement Sommative) ou Je ne sais pas ;
- (2) une estimation du niveau des indicateurs, parmi 3 choix : Faible (F), Moyen (M) ou Élevé (E) ;
- (3) une valeur numérique pour chaque indicateur entre 0 et 1.

## 5 Calcul des Indicateurs

### 5.1 Comparaison de Séquences de Textes

Afin de construire nos indicateurs, nous examinons d'abord l'évolution d'un texte afin de la quantifier. Pour ce faire, nous comparons des séquences de textes par paires, en utilisant la méthode dite *Sequence Matcher* dérivant de l'algorithme *Gestalt Pattern Matching* [19]<sup>3</sup>. Cette méthode examine l'évolution d'un texte entre deux versions. Cela consiste tout d'abord à repérer la plus longue (en termes de nombre de caractères) séquence contiguë d'appariement qui ne contient aucun élément dit "inutile", tels que les lignes vides ou les espaces blancs. Ensuite, cette opération est répétée pour les séquences se trouvant à gauche et à droite de cette séquence contiguë. Enfin, afin de nommer les changements dans le texte, une étiquette est associée à chaque séquence : *Equal* (les séquences sont égales), *Insert* (la séquence est insérée), *Delete* (la séquence est supprimée), *Replace* (la séquence est remplacée).

### 5.2 Matrice de Contribution

Nous désignons par *matrice de contribution*, une matrice mathématique qui fournit pour chaque mot d'un texte final, le niveau de contribution de chacun des apprenants ayant co-écrit ce texte. Dans cette matrice, les lignes représentent les apprenants contributeurs (co-auteurs) et les colonnes représentent les mots constituant le texte final. C'est une matrice  $K \times M$ , où  $K$  est le nombre d'auteurs et  $M$  le nombre de mots co-écrits par les  $K$  auteurs. Dans cette matrice, le niveau de contribution d'un auteur  $i$  à un mot  $l$  dans une

<sup>3</sup> *Bibliothèque DiffLib* <https://github.com/python/cpython/blob/main/Lib/difflib.py>

phrase  $j$  est défini par  $x_{i,j,l} \in [0, 1]$ , où  $i \in [1, K]$ ,  $j \in [1, N]$  et  $l \in [1, n_j]$ ,  $N$  est le nombre de phrases composant le texte, et  $n_j$  désigne le nombre de mots dans la phrase  $j$ . Il est égal à 1 lorsque le mot est entièrement écrit par cet auteur, et à 0, lorsque l'auteur ne contribue pas à l'écriture de ce mot. La matrice de contribution ne prend pas en compte les mots supprimés. Le total des contributions des auteurs à un mot est égal à 1.

### 5.3 Balance de Contribution

La balance de contribution est un indicateur qui mesure combien les contributions d'écriture collaborative des apprenants sont égales (bien équilibrées) ou non en termes de mots constituant le texte. Plus la valeur de cet indicateur est proche de 1, plus les contributions des apprenants sont égales ou équilibrées. À l'inverse, plus il est proche de 0, moins les contributions des apprenants sont égales ou équilibrées. Cet indicateur est basé sur la variance des contributions moyennes des auteurs, reflétant la distance entre la contribution moyenne de chaque auteur ( $\bar{x}_{i,..}$ ) et les contributions moyennes de l'ensemble des auteurs aux mots composant le texte. Notons que, si les auteurs contribuent de manière équilibrée, alors leurs scores sont proches de la moyenne et ils contribuent de manière tout à fait égale. De plus, afin de pénaliser le cas où le texte entier est écrit par un seul auteur, cette mesure est normalisée. Enfin, par souci de facilité d'interprétation pour les enseignants, cet indicateur est donné par 1 moins une valeur de dispersion, ce qui donne l'indicateur balance de contribution comme suit

$$e(X) = 1 - \frac{K}{K-1} \sum_{i=1}^K \left( \bar{x}_{i,..} - \frac{1}{K} \right)^2. \quad (1)$$

Notons que,  $\sum_{i=1}^K \bar{x}_{i,..} = 1$  et que la moyenne des contributions moyennes de l'ensemble des auteurs est donnée par

$$\bar{\bar{x}}_{,..} = \frac{1}{K} \sum_{i=1}^K \bar{x}_{i,..} = \frac{1}{K}.$$

Cet indicateur (1) atteint sa valeur maximale 1, lorsque la valeur de la dispersion soustraite à 1 atteint sa valeur minimale 0, autrement dit, lorsque tous les auteurs contribuent de manière égale (ou équilibrée) au texte, c'est-à-dire lorsque la contribution moyenne de chaque auteur  $\bar{x}_{i,..}$  est égale à  $1/K$ . Elle atteint sa valeur minimale 0 lorsqu'un auteur écrit seul la totalité du texte.

### 5.4 La Co-écriture

À l'instar de la balance de contribution, mesurant les contributions des auteurs en terme de mots à un texte, l'indicateur de co-écriture mesure les contributions des auteurs au niveau des phrases. Ainsi, la construction de cet indicateur nécessite en amont le découpage du texte en phrases. Pour ce faire, nous utilisons une approche de segmentation de phrases, basée sur des règles (heuristiques) pour sa simplicité de mise en œuvre [20]. Nous utilisons un outil de détection des limites de phrases basé sur un ensemble de

règles appelé *Golden Rule Set*<sup>4</sup>, conçues à la main permettant de déterminer les limites de phrases, telles que les ponctuations.

Grâce à la matrice de contribution, nous calculons pour chaque auteur sa contribution moyenne à l'ensemble des phrases. Nous pénalisons le cas où un auteur écrit seul une phrase, et nous calculons pour une phrase  $j$ , 1 moins une valeur de dispersion, comme suit

$$e_j(X) = 1 - \frac{K}{K-1} \sum_{i=1}^K \left( \bar{x}_{i,j,\cdot} - \frac{1}{K} \right)^2.$$

Ainsi, l'indicateur de co-écriture est donné par

$$c(X) = \sum_{j=1}^N p_j e_j(X) \quad \text{où} \quad p_j = \frac{n_j}{M} \quad (2)$$

est le poids de la phrase  $j$ . Lorsque toutes les phrases sont écrites par un seul auteur, l'indicateur de co-écriture est égal à 0. Il est égal à 1 lorsque toutes les phrases sont co-écrites de manière équilibrée par l'ensemble des auteurs.

Les deux métriques (1) et (2) permettent de répondre à la QR1. Par ailleurs, les deux indicateurs possèdent une propriété découlant de ces métriques. En effet, l'indicateur de co-écriture  $c(X)$  dans (2) est inférieur ou égal à l'indicateur balance de contribution  $e(X)$  dans (1) [9].

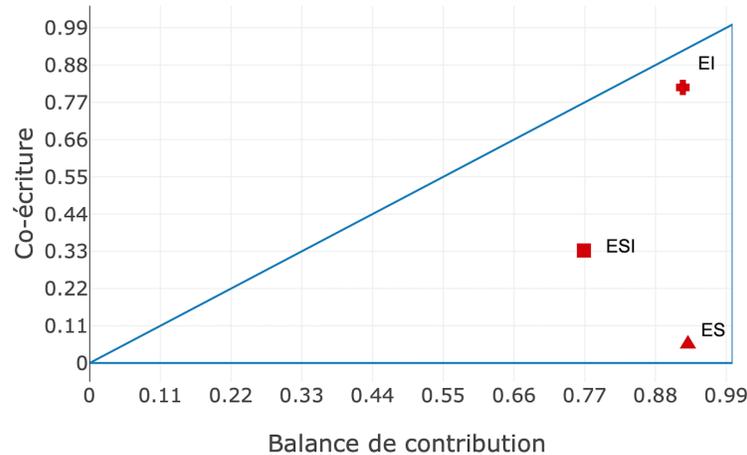
### 5.5 Relation des Indicateurs aux Stratégies d'Écriture Collaborative

Ces indicateurs ont des valeurs continues comprises entre 0 et 1. Par conséquent, nous pouvons représenter les documents texte co-écrits dans un plan à deux dimensions où l'axe des abscisses correspond à la balance de contribution et l'axe des ordonnées correspond à la co-écriture (Figure 1). De plus, étant donné que l'indicateur de co-écriture est toujours inférieur ou égal à la balance de contribution, l'espace de représentation des documents dans le plan 2D est réduit à un demi plan.

Il est alors possible de distinguer les documents écrits selon des stratégies sommatives (ES, PS) ou intégratives (EI, PI) de ceux écrits selon une stratégie mixte (ESI), grâce à leur emplacement dans ce plan 2D. Cela nous permet de répondre à la QR2. En effet, lorsque la co-écriture est faible, la stratégie est sommative. Par conséquent, les documents écrits selon des stratégies sommatives sont représentés près de l'axe des  $x$ . Lorsque les stratégies sont intégratives, les documents sont représentés près de la diagonale du plan 2D et les valeurs des deux indicateurs sont proches. Les documents rédigés selon une stratégie mixte sont représentés au milieu de l'espace entre l'axe des  $x$  et la diagonale.

Nous illustrons dans la figure 1 trois exemples de documents, un premier écrit avec une stratégie intégrative (EI) ayant une balance de contribution égale à 0.92 et de co-écriture égale à 0.79, un second document écrit avec une stratégie sommative (ES) ayant comme coordonnées (0.92, 0.05) et un troisième écrit selon une stratégie mixte (ESI) dont les coordonnées sont (0.77, 0.33).

<sup>4</sup>[https://github.com/diasks2/pragmatic\\_segmenter](https://github.com/diasks2/pragmatic_segmenter)



**Fig. 1.** Plan 2D montrant le lien entre les indicateurs et les stratégies d'écriture collaborative de documents texte.

## 6 Résultats et Discussion

Les résultats du questionnaire nous ont permis d'évaluer les interprétations des stratégies et des indicateurs par les enseignants (QR3). Nous avons comparé les valeurs et les niveaux des indicateurs estimés par les enseignants aux valeurs machine, i.e. calculées à l'aide des formules (1) et (2). Le niveau d'un indicateur est faible, si sa valeur se situe dans  $[0, 1/3[$ , moyen si sa valeur se situe dans  $[1/3, 2/3[$  et élevé si sa valeur se situe dans  $[2/3, 1]$ .

Le tableau 1 compare pour chaque LabDoc et pour chaque stratégie le nombre de réponses reçues (en gras les LabDocs mal classés). On remarque que 9/12 LabDocs sont bien classés avec une majorité de bonnes réponses variant entre 8 et 14. Les LabDocs 1 et 6 sont mal classés, mais présentent respectivement 6 et 5 bonnes réponses. Ceci est probablement dû aux formules mathématiques présentes dans le LabDoc 1, pouvant influencer la perception des enseignants (les formules mathématiques n'étant pas prises en compte à ce stade), et à la difficulté pour les enseignants de percevoir la stratégie mixte du LabDoc 6. Pour le LabDoc 12, les réponses des enseignants sont plus dispersées par rapport aux LabDocs 1 et 6, mais avec 5 bonnes réponses et 1 réponse "Je ne sais pas". Le LabDoc 12 présente un niveau de collaboration très faible, pouvant conduire à une mauvaise interprétation de la stratégie par les enseignants.

Les enseignants ont donné une estimation des indicateurs, en choisissant un niveau parmi les 3 proposés (faible, moyen et élevé), et une valeur numérique entre 0 et 1. Le tableau 2 fournit pour chaque LabDoc le nombre d'apprenants ayant contribué à sa rédaction, le nombre de réponses des enseignants pour chaque niveau d'indicateur (faible, moyen et élevé), le niveau de chaque indicateur, ainsi que sa valeur machine. Nous avons également calculé l'écart entre la valeur machine (VM) d'un indicateur

**Table 1.** Vérification des stratégies perçues par les enseignants. En gras on distingue les cas où la majorité des réponses des enseignants ne correspond pas à la stratégies réelle d'un LabDoc.

LabDoc	Stratégie	ES	PS	ESI	PI	EI	Je ne sais pas
<b>1</b>	<b>EI</b>	<b>0</b>	<b>2</b>	<b>4</b>	<b>3</b>	<b>6</b>	<b>0</b>
2	ES	11	2	1	1	0	0
3	EI	0	1	1	1	12	0
4	ESI	0	1	8	5	1	0
5	EI	2	0	0	2	11	0
<b>6</b>	<b>ESI</b>	<b>0</b>	<b>8</b>	<b>5</b>	<b>2</b>	<b>0</b>	<b>0</b>
7	ES	12	1	0	0	2	0
8	PI	0	0	4	8	3	0
9	PS	0	8	5	1	0	1
10	ES	14	0	1	0	0	0
11	ES	14	1	0	0	0	0
<b>12</b>	<b>EI</b>	<b>1</b>	<b>4</b>	<b>2</b>	<b>2</b>	<b>5</b>	<b>1</b>

$\theta \in [0, 1]$  et l'estimation des enseignants, en utilisant l'Écart Quadratique Moyen (EQM)

$$EQM = \sqrt{\frac{\sum_1^n (\hat{\theta}_i - \theta)^2}{n}}$$

où, pour  $i = 1, \dots, n$ ,  $\hat{\theta}_i \in [0, 1]$  est l'estimation d'un enseignant et  $n = 15$  est le nombre d'enseignants ayant répondu au questionnaire.

Les résultats du tableau 2 montrent que 8/12 LabDocs sont bien classés par les enseignants concernant les niveaux de l'indicateur balance de contribution et 10/12 concernant les niveaux de l'indicateur de co-écriture. Nous remarquons que l'EQM prend des valeurs dans l'intervalle  $[0.07, 0.39]$  concernant la balance de contribution et dans l'intervalle  $[0.04, 0.42]$  concernant la co-écriture. Les résultats montrent que lorsque les enseignants parviennent à percevoir un niveau correct pour un indicateur ils parviennent aussi à bien estimer sa valeur. De plus, concernant la balance de contribution, les LabDocs pour lesquels le niveau n'est pas correctement estimé, sont ceux écrits par plus de deux apprenants sauf pour le LabDoc 7. Concernant la co-écriture, le niveau n'est pas correctement estimé pour le LabDoc 1 (probablement à cause de la présence de formules mathématiques) et le LabDoc 5 (probablement à cause de la difficulté de percevoir correctement la co-écriture sur un LabDoc rédigé par plus de deux auteurs, et selon une stratégie intégrative).

A partir de ces résultats, et en réponse à la QR3, nous pouvons déduire que les indicateurs et les stratégies sont bien interprétés par les enseignants dans une certaine mesure. Cependant, les stratégies des LabDocs comprenant des formules mathématiques, ou rédigés selon des stratégies mixtes ou présentant un très faible niveau de collaboration peuvent être difficiles à percevoir. En ce qui concerne les indicateurs, des difficultés dans leur estimation sont également observées pour les LabDocs écrits par plus de 2 co-auteurs.

**Table 2.** Comparaison des niveaux d'indicateurs perçus Faible (F), Moyen (M) et Élevé (E) et des valeurs estimées par les enseignants avec les niveaux (NM) et les valeurs machines (VM). En gras on distingue les cas où la majorité des réponses des enseignants ne correspondent pas aux résultats des mesures calculées par la machine. NB désigne le nombre de co-auteurs d'un LabDoc et EQM l'écart quadratique moyen.

LabDoc	NB	Balance de contribution						Co-écriture					
		F	M	E	NM	VM	EQM	F	M	E	NM	VM	EQM
1	2	0	8	7	M	0.49	0.32	<b>0</b>	<b>4</b>	<b>11</b>	<b>M</b>	<b>0.41</b>	<b>0.42</b>
2	2	0	6	9	E	0.87	0.15	13	1	1	F	0.19	0.11
3	2	1	2	12	E	0.84	0.24	1	0	14	E	0.76	0.28
4	2	0	1	14	E	0.99	0.21	3	9	3	M	0.41	0.22
5	3	<b>15</b>	<b>0</b>	<b>0</b>	<b>M</b>	<b>0.49</b>	<b>0.36</b>	<b>12</b>	<b>1</b>	<b>2</b>	<b>M</b>	<b>0.44</b>	<b>0.35</b>
6	2	0	1	14	E	0.99	0.19	7	7	1	M	0.50	0.2
7	2	<b>13</b>	<b>2</b>	<b>0</b>	<b>M</b>	<b>0.51</b>	<b>0.3</b>	14	1	0	F	0.07	0.09
8	2	0	7	8	E	0.99	0.3	0	6	9	E	0.74	0.21
9	4	<b>13</b>	<b>2</b>	<b>0</b>	<b>M</b>	<b>0.48</b>	<b>0.27</b>	13	2	0	F	0.17	0.12
10	3	<b>1</b>	<b>11</b>	<b>3</b>	<b>E</b>	<b>0.92</b>	<b>0.39</b>	13	2	0	F	0.09	0.1
11	2	15	0	0	F	0.04	0.04	15	0	0	F	0.29	0.19
12	2	15	0	0	F	0.14	0.07	14	1	0	F	0.09	0.09

## 7 Conclusion

Ce travail contribue à la recherche sur l'écriture collaborative, en proposant deux nouveaux indicateurs. Le premier permet de mesurer l'égalité des contributions en terme de mots composant un texte (balance des contributions), et le second mesure les interactions entre apprenants en termes des contributions au niveau des phrases (co-écriture).

Une deuxième contribution de l'article consiste à identifier les stratégies d'écriture collaborative sommative et intégrative qui sont bien définies dans la littérature et utiles dans un contexte pédagogique, par exemple pour distinguer les travaux coopératifs (sommatifs) des travaux collaboratifs (intégratifs) [17]. Nous avons montré que ces stratégies d'écriture peuvent être dérivées des deux indicateurs proposés.

Une troisième contribution de cet article consiste en l'évaluation de l'interprétation des indicateurs et des stratégies proposées. Nous avons utilisé des documents texte édités dans la plateforme *LabNbook* et issus d'un contexte réel. Nous avons réalisé une évaluation par questionnaire auprès d'enseignants qui ne sont pas des utilisateurs experts de *LabNbook*. Les résultats du questionnaire ont permis de comparer la perception des indicateurs par les enseignants avec les résultats des calculs automatiques, et de vérifier les stratégies perçues. Les résultats ont montré que les enseignants ont une bonne interprétation des indicateurs et des stratégies. En effet, les enseignants étaient capables d'estimer et de percevoir correctement les indicateurs et les stratégies de co-écriture de la plupart des documents.

Enfin, nous concluons que les nouvelles métriques proposées pour la balance de contribution et la co-écriture peuvent aider à évaluer le travail collaboratif d'apprenants, en caractérisant le niveau de collaboration et les stratégies d'écriture collaborative. En

perspective, il serait intéressant d'étudier comment permettre aux enseignants d'exploiter ces indicateurs à partir de visualisations appropriées, et d'évaluer leur acceptabilité, utilisabilité et utilité. De plus, ce travail n'est pas sans limites. Il serait pertinent d'améliorer notre approche d'analyse de texte, afin de prendre en compte les formules mathématiques. En effet, notre méthode de comparaison de séquences de textes peut produire des résultats erronés lorsque les documents texte comportent des formules mathématiques. Une amélioration possible est de repérer les formules mathématiques en utilisant les techniques d'apprentissage automatique, comme les Chaînes de Markov Cachées.

**Remerciement.** Le projet est co-financé par la Région Bretagne dans le cadre de son programme Stratégie d'Attractivité Durable (SAD) et le Conseil départemental du Finistère dans le cadre de son dispositif Aide aux Programmes de Recherche Emergents (APRE) (2021-2023). Les auteurs tiennent à remercier l'ensemble des enseignants ayant contribué à cette étude.

## Références

1. Abrams, Z. : Exploring collaboratively written 12 texts among first-year learners of german in google docs. *Computer Assisted Language Learning* **29**(8), 1259–1270 (2016)
2. Avenier, M.J., Thomas, C. : Finding one's way around various methodological guidelines for doing rigorous case studies : A comparison of four epistemological frameworks. *Systemes d'information management* **20**(1), 61–98 (2015)
3. Baker, M. : A model for negotiation in teaching-learning dialogues. *Journal of Interactive Learning Research* **5**(2), 199 (1994)
4. Chen, J., Wang, M., Kirschner, P.A., Tsai, C.C. : The role of collaboration, computer use, learning environments, and supporting strategies in cscl : A meta-analysis. *Review of Educational Research* **88**(6), 799–843 (2018)
5. Cho, H. : Synchronous web-based collaborative writing : Factors mediating interaction among second-language writers. *Journal of Second Language Writing* **36**, 37–51 (2017)
6. d'Ham, C., Wajeman, C., Girault, I., Marzin-Janvier, P. : LabNbook, plateforme numérique support des pédagogies actives et collaboratives en sciences expérimentales. In : Broisin, J., Sanchez, E., Yessad, A., Chenevotot, F. (eds.) *EIAH 2019 : Environnement Informatiques pour l'Apprentissage Humain*. pp. 49–60. Actes de la 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, Paris, France (Jun 2019)
7. Dillenbourg, P. : What do you mean by collaborative learning?, p. 1–19. Elsevier, Oxford (1999)
8. Dillenbourg, P., Michael, B. : Negotiation spaces in human-computer collaborative learning. In : *Proceedings of the International Conference on Cooperative Systems*. p. 12–14 (1996)
9. Haddouche, M.A., Djelil, F., Hoffmann, C., Mandran, N., d'Ham Cédric. : Proposal of indicators for measuring collaborative writing in a digital learning environment. *15th International Conference on Computer Supported Education* **2**, 495–502 (2023)
10. Hoffmann, C., Mandran, N., d'Ham, C., Rebaudo, S., Haddouche, M.A. : Development of actionable insights for regulating students' collaborative writing of scientific texts. In : *Educating for a New Future : Making Sense of Technology-Enhanced Learning Adoption*. p. 534–541. *Lecture Notes in Computer Science*, Springer International Publishing, Cham (2022)

11. Larsen-Ledet, I., Korsgaard, H. : Territorial functioning in collaborative writing. *Computer Supported Cooperative Work (CSCW)* **28**(3), 391–433 (Jun 2019)
12. Li, M., Kim, D. : One wiki, two groups : Dynamic interactions across esl collaborative writing tasks. *Journal of second language writing* **31**, 25–42 (2016)
13. Li, M., Zhu, W. : Explaining dynamic interactions in wiki-based collaborative writing. *Language Learning & Technology* **21**(2), 96–120 (2016)
14. Lowry, P.B., Curtis, A., Lowry, M.R. : Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice. *The Journal of Business Communication* (1973) **41**(1), 66–99 (2004)
15. Mandran, N., Vermeulen, M., Prior, E. : Thedre's framework : Empowering phd candidates to efficiently implement design-based research. *Education and Information Technologies* pp. 1–24 (2022)
16. Olson, J.S., Wang, D., Olson, G.M., Zhang, J. : How people write together now : Beginning the investigation with advanced undergraduates in a project course. *ACM Transactions on Computer-Human Interaction* **24**(1), 4 :1–4 :40 (Mar 2017)
17. Onrubia, J., Engel, A. : Strategies for collaborative writing and phases of knowledge construction in cscl environments. *Computers & Education* **53**(4), 1256–1265 (2009)
18. Posner, I.R., Baecker, R.M. : How people write together (groupware). In : *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*. vol. 4, p. 127–138. IEEE (1992)
19. Ratcliff, J.W., Metzener, D.E. : Pattern-matching-the gestalt approach. *Dr Dobbs Journal* **13**(7), 46 (1988)
20. Sadvilkar, N., Neumann, M. : PySBD : Pragmatic sentence boundary disambiguation. In : *Proceedings of Second Workshop for NLP Open Source Software (NLP-OSS)*. pp. 110–114. Association for Computational Linguistics, Online (Nov 2020)
21. Storch, N. : *Collaborative Writing in L2 Classrooms*. Multilingual Matters, Bristol, Blue Ridge Summit (2013)
22. Sundgren, M., Jaldemark, J. : Visualizing online collaborative writing strategies in higher education group assignments. *The International Journal of Information and Learning Technology* **37**(5), 351–373 (Aug 2020)
23. Wang, D., Tan, H., Lu, T. : Why users do not want to write together when they are writing together : Users' rationales for today's collaborative writing practices. *Proceedings of the ACM on Human-Computer Interaction* **1**(CSCW), 1–18 (2017)
24. Wang, F., Hannafin, M.J. : Design-based research and technology-enhanced learning environments. *Educational technology research and development* **53**(4), 5–23 (2005)
25. Zhang, M., Chen, W. : Assessing collaborative writing in the digital age : An exploratory study. *Journal of Second Language Writing* **57**, 100868 (sep 2022)
26. Zhang, M., Gibbons, J., Li, M. : Computer-mediated collaborative writing in l2 classrooms : A systematic review. *Journal of Second Language Writing* **54**, 100854 (2021)