In the context of solely trace-based indicators, we address in this paper the following research questions:

RQ 1: What are the indicators (metrics) allowing to measure collaborative writing ?

RQ 2: What are the collaborative writing strategies that can be deduced from these indicators ?

RQ 3: To what extent these indicators and strategies are interpretable by teachers ? (How close is the relationship between these measures and their interpretation).

**1 Introduction**

peut-être ici, il faudrait annoncer les deux indicateurs construits et leur propriété (intervalle 0-1) afin qu’on puisse s’y référer dans la partie Data collection

**2 Theoretical Background**

Introduire ici les deux stratégies, summative et integrative, en se basant sur la réf Onrubia & Engel ;

-> permettent de faire le lien avec la distinction entre coopération et collaboration

**3 Research Method**

3.1 Context

3.2 Data collection

3.2.1 Focus Group

3.2.2 Questionnaires

We used a questionnaire to measure to what extent the indicators are interpretable by teachers who are not expert users of Labnbook (RQ3). Our objective is to verify whether the definitions of the indicators make sense. In the LabNbook database, we selected 12 LabDoc co-written by 2 to 4 students in real conditions. The writing contributions of different students were highlighted. These Labdocs are chosen with different writing strategies: 4 Labdocs written with an integrative strategy (Entirely Integrative - EI), 4 Labdocs with a summative strategy (Entirely Summative - ES), and 4 Labdocs with intermediate strategies: part of the text was written with a summative strategy and part with an integrative strategy. For the latter, the text of 2 LabDocs is roughly written half/half (Between summative and integrative - BSI)), for one mostly summatively (Rather summative - RS), and the last is mostly integrative (Rather integrative - RI).

Teachers are asked to read the 12 LabDocs. After reading each of them, they indicate whether they perceive collaboration in the written documents (yes, no, perhaps) and argue their responses. In addition, for each Labdoc, we ask the teachers to indicate :

(1) the writing strategy between 6 choices: EI, RI, BSI, RS, ES or I don’t know

(2) a rough estimate for the indicators, using 3 levels: Low (L), Medium (M) or High (H) ; *? donner les intervalles (0 - 0.33 ; 0.33 -0.66 ; 0.66 -1) ?*

(3) a numerical value for each indicator between 0 and 1.

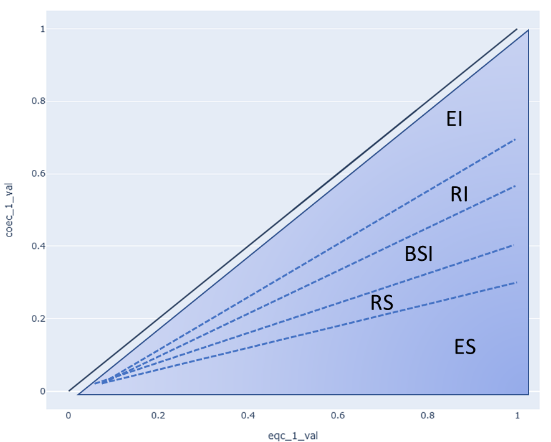
The questionnaire ended with questions on teaching practices and student monitoring. The questionnaire results are interpreted in Section 5.

**4 Indicator construction**

4.5

~~Note that, it can be more convenient for a teacher to get a qualitative description of the indicators in-  
stead of a value. It is also more practice for evaluating the indicators through a questionnaire. Hence, we  
split the indicators interval of values [0, 1] on threelevels as follows: Low for values in ]0, 1/3]; Medium  
for values in [1/3, 2/3[; High for values in [2/3, 1].~~Combining the two indicators and using property (10), we are able to distinguish CW strategies, e.g. the sequential summative construction and the sequential integrating construction defined above. This leads to represent LabDocs in the  
triangle in Figure 1. For instance, in a LabDoc with balance of contribution and co-writing scores equals 0.8 and 0.65 respectively, our approach suggest that the CW strategie adopted by students in their CW process is entirely integrative (EI)

à discuter si on met la figure :



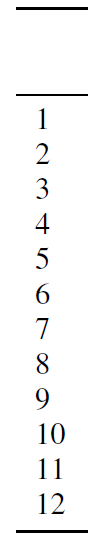
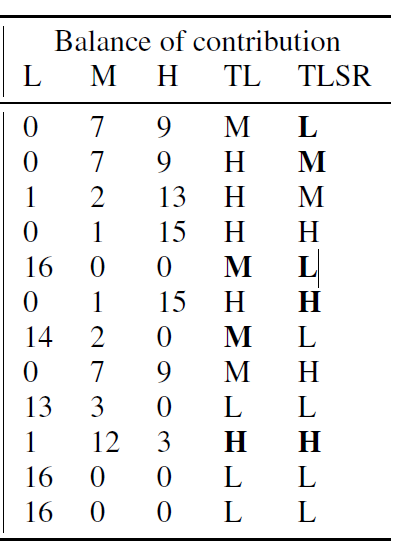
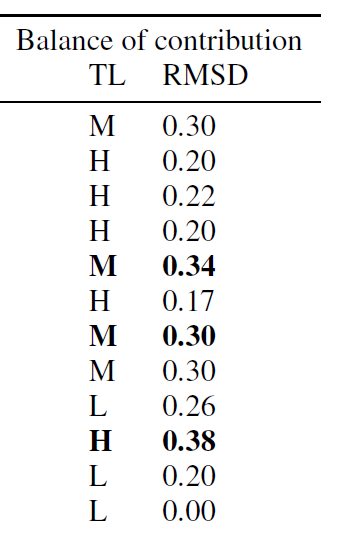
**5 Résultats de l’évaluation des indicateurs auprès des enseignants**

**5.1 Niveaux et valeurs numériques des deux indicateurs**

**Balance de contribution**

|  |  |  |
| --- | --- | --- |
| **LD** | **Indicateur / machine** | **Réponses des enseignants** |

|  |  |
| --- | --- |
| **MV** | Level |
| 0,49 | M |
| 0,87 | H |
| 0,84 | H |
| 0,99 | H |
| 0,49 | **M** |
| 0,99 | H |
| 0,51 | **M** |
| 0,99 | H |
| 0,48 | **M** |
| 0,92 | **H** |
| 0,30 | L |
| 0,14 | L |

estimation du niveau

résultat : dans le tableau

🡪 8/12 bien classés

Estimation de la valeur numérique

les MV : BC varie entre 0.14 et 0.99

résultat : rmsd (vérifier que la signification du RMSD est bien comprise par Cédric et Fahima)

🡪 à commenter en fonction des résultats : donner la fourchette 0.07 … 0.38 (à remettre à jour) - dire que cela traduit une dispersion assez importante

**Co-écriture**

estimation du niveau de l’indicateur CW par les enseignants entre L, M et H

résultat : dans le tableau

🡪 10/12 bien classés

estimation chiffrée

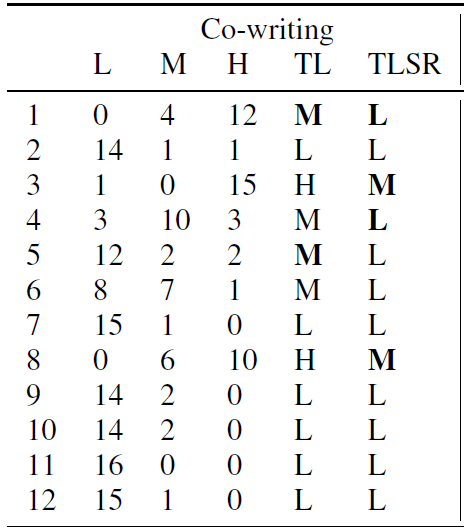
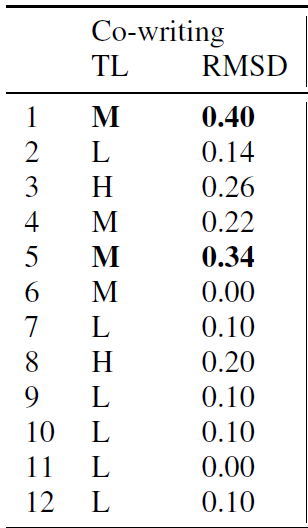
les MV : CW entre 0.04 et 0.77

résultat : rmsd

🡪 à commenter en fonction des résultats : donner la fourchette 0.09 et 0.39

- dire que cela traduit une dispersion assez importante

|  |  |  |
| --- | --- | --- |
| **LD** | **Indicateur** | **Réponses des enseignants** |



|  |  |  |
| --- | --- | --- |
| **LD** | **MV** | Level |
| 1 | 0,41 | **M** |
| 2 | 0,11 | L |
| 3 | 0,77 | H |
| 4 | 0,41 | M |
| 5 | 0,44 | **M** |
| 6 | 0,51 | M |
| 7 | 0,07 | L |
| 8 | 0,74 | H |
| 9 | 0,17 | L |
| 10 | 0,09 | L |
| 11 | 0,04 | L |
| 12 | 0,10 | L |

*à discuter si on ajoute le Rayon de chaque Labdoc ?*

**5.2 Stratégies d’écriture collaborative**

(tableau 5 mais commencer par les dernières colonnes : stratégie et tendance, cf. ci-dessous)

Question 1 du questionnaire : (vise à répondre à la question si les enseignants classent les LabDocs dans les mêmes stratégies que les experts qui ont conçu les indicateurs)

Résultat :

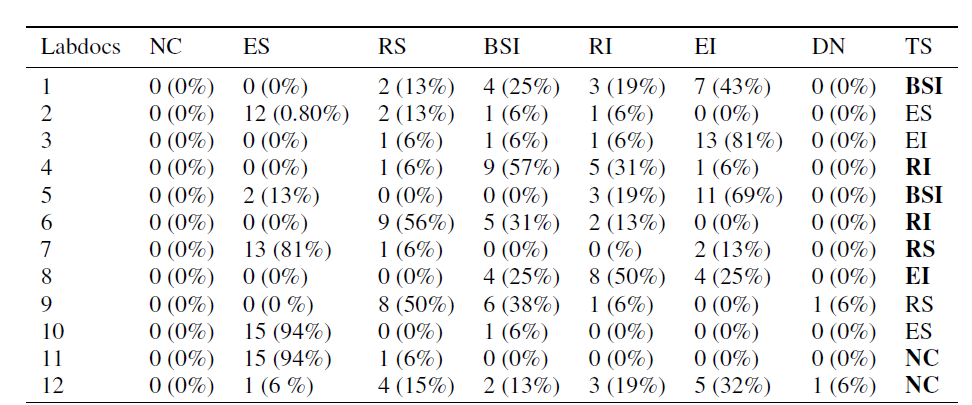
pour 9 LabDocs la majorité des enseignants les a bien classés (entre 50% et 94%) ;

exceptions à discuter en détail :

LD1 : 43% EI ok et 19% dans RI

LD 6 : 31% - BSI - ok et 56% RS

LD12 : 32% EI - ok, 19 % RI et 13% BSI (classement étalé)





**6. Discussion**

*à voir avec Nadine :*

*- est-ce qu’on distingue la partie Résultats (« brut ») en section 5 et leur discussion/ interprétation en section 6 ?*

*- est-ce qu’on inclut la discussion des RQ1 et RQ2 en section 6 ou on la met dans la partie Conclusion ?*

*à voir où le placer dans l’article :*

[ RQ1 : What are the indicators (metrics) allowing to measure collaborative writing ?

*à discuter …*

- notre choix des deux indicateurs se base sur la littérature

- faut-il y revenir ici ? -> les deux indicateurs séparément donnent des infos pertinentes pour caractériser le CW (référence aux concepts de ‘equality’ et ‘mutuality’)

RQ 2: What are the collaborative writing strategies that can be deduced from these indicators ?

- nous avons exploré l’exemple de deux classes de stratégie importantes : summative et intégrative

- si on veut élargir à d’autres stratégies, il faut donner des exemples ;

- sinon on pourrait aussi réfléchir à ce qui manque comme info pour distinguer d’autre stratégies (parler des limites…) ]

--------------------

RQ 3: To what extent these indicators and strategies are interpretable by teachers ? (How close is the relationship between these measures and their interpretation).

*Commenter les résultats de la section 5*

BC : estimations du niveau ok ? (8/12)

donner des valeurs numériques est plus difficiles, surtout si

- plus de 2 étudiants (LD5 et LD10) (*il faut donner l’info qqpart que les LabDocs écrits par plus de 2 étudiants sont 5, 9, 10*)

- valeurs intermédiaires, autour de 0.5 (LD7 et LD 9)

CW : idée également perçu par les enseignants

estimation du niveau OK (10/12)

Mais :

- limitation à BC (qui vient du fait qu’on tient également compte de l’équilibre de contribution, au niveau d’une phrase) n’est pas bien perçu (aussi dû au fait que ce n’était pas suffisamment expliqué dans le questionnaire) ; exemples LD 1, LD5

🡪 les enseignants ont pensé à deux indicateurs indépendants

au niveau des stratégies :

- les enseignants reconnaissent les stratégies

* si on décide d’ajouter les rayons de chaque LabDoc, il faudrait le commenter aussi :

🡪 rmsd assez important -> discuter les diverses raisons : contrainte BC < CW non-perçue ; écriture à plus de 2 auteurs plus difficile à interpréter ; la présence de formules dans les LabDocs ; confusion possible entre la fin de la phrase par le ‘. ‘ et une nouvelle ligne dans nos LD exemples (est-ce que la phrase en intro du doc a été lue ?) ; etc.

Anis. :

Si on cite les LD pour on devrait les rendre accéssible dans le papier. Le lecteur doit les voir. (matériel supplémentaire ?)

**7. Conclusions**

quand ils s’agit de construire des indicateurs à fournir aux enseignants :

- des indicateurs descriptifs qui sont faciles à comprendre/interpréter (CW trop complexe ?)

- si on propose plusieurs, ils devraient être indépendants

🡪 doivent être « actionable » c’est-à-dire permettre d’engager des actions concrète dans l’enseignement, l’accompagnement des étudiants

quand il s’agit de créer des indicateurs pour un traitement automatique de traces, p.ex. analyser la qualité d’un travail collaboratif et renvoyer une information à ce sujet aux enseignants (alerte, curseur, etc.) afin qu’ils engagent la discussion avec les étudiants :

- CW et BC sont de bons candidats, car ils caractérisent bien le degré et le type de collaboration

- ils pourraient être utilisé p.ex. comme paramètres d’entrée dans des modèles de machine learning (cf p.ex. papier Chejara, 2020)

-> à explorer dans le futur ?

Anis :

Je rajouterais:

la limite des deux indicateurs quant à la présence des formules.

Dire qu’une perspective possible est d’entrainer un model ML pour extraire les formules et les traiter séparémeent.

Fahima : problème d'effet de seuil

Annexe :

*[juste pour nous (pas à mettre dans l’article), voici comment les 12 LabDocs se repartissent dans le plan CW-BC]*

