



HAL
open science

DRAXO, Un cadre pour l'évaluation formative par les pairs

Franck Silvestre

► **To cite this version:**

Franck Silvestre. DRAXO, Un cadre pour l'évaluation formative par les pairs. Quatrième Journées des INSPE d'Occitanie, Jun 2022, Agde (34), France. hal-04186725

HAL Id: hal-04186725

<https://hal.science/hal-04186725>

Submitted on 24 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.



DRA XO

Un cadre pour l'évaluation formative par les pairs

Franck Silvestre - franck.silvestre@irit.fr

06/2022

Contexte

- Processus d'évaluation formative
- Questions de compréhension
- Promotion de l'argumentation écrite et de l'instruction par les pairs
- Support technologique : Elaastic

L'enseignant pose
une question



▼ **Fonctions polynomiales** [QUESTION À CHOIX EXCLUSIF]

$$g(x) = ax^2 + 24$$

Dans la fonction g , a est une constante et $g(4) = 8$.

Quelle est la valeur de $g(-4)$?

1. $g(-4)$ est égal à 8
2. $g(-4)$ est égal à 0
3. $g(-4)$ est égal à -1
4. $g(-4)$ est égal à -8

L'enseignant pose
une question



Les apprenants répondent
à la question



Votre réponse : 1 2 3 4

Réponse textuelle

B *I* U | I_x       | ”    

Etant donné que $g(4)$ est positif, alors $g(-4)$ renvoie le même résultat mais négatif.

Votre degré de confiance

Pas du tout confiant(e)

Pas vraiment confiant(e)

Confiant(e)

Tout à fait confiant(e)

Soumettre

L'enseignant pose une question



Les apprenants répondent à la question



Les apprenants confrontent leurs points de vue



Les apprenants répondent à la question



Confrontation de points de vue

Ci-dessous sont présentées une ou plusieurs réponses alternatives. Merci de bien vouloir indiquer dans quelle mesure vous êtes d'accord avec ces réponses.

Choix [4]

Si $g(4)$ retourne 8

Votre évaluation :

Pas du tout d'accord

Tout à fait d'accord

Je ne me prononce pas 1 2 3 4 5

Choix [1]

x^2 retourne la même valeur pour n'importe quel nombre et son opposé. Par conséquent $g(4) = g(-4) = 8$

Votre évaluation :

Je ne me prononce pas 1 2 3 4 5

Vous disposez d'une deuxième chance pour changer votre réponse et votre degré de confiance.

Votre réponse : 1 2 3 4

L'enseignant pose une question



Les apprenants répondent à la question



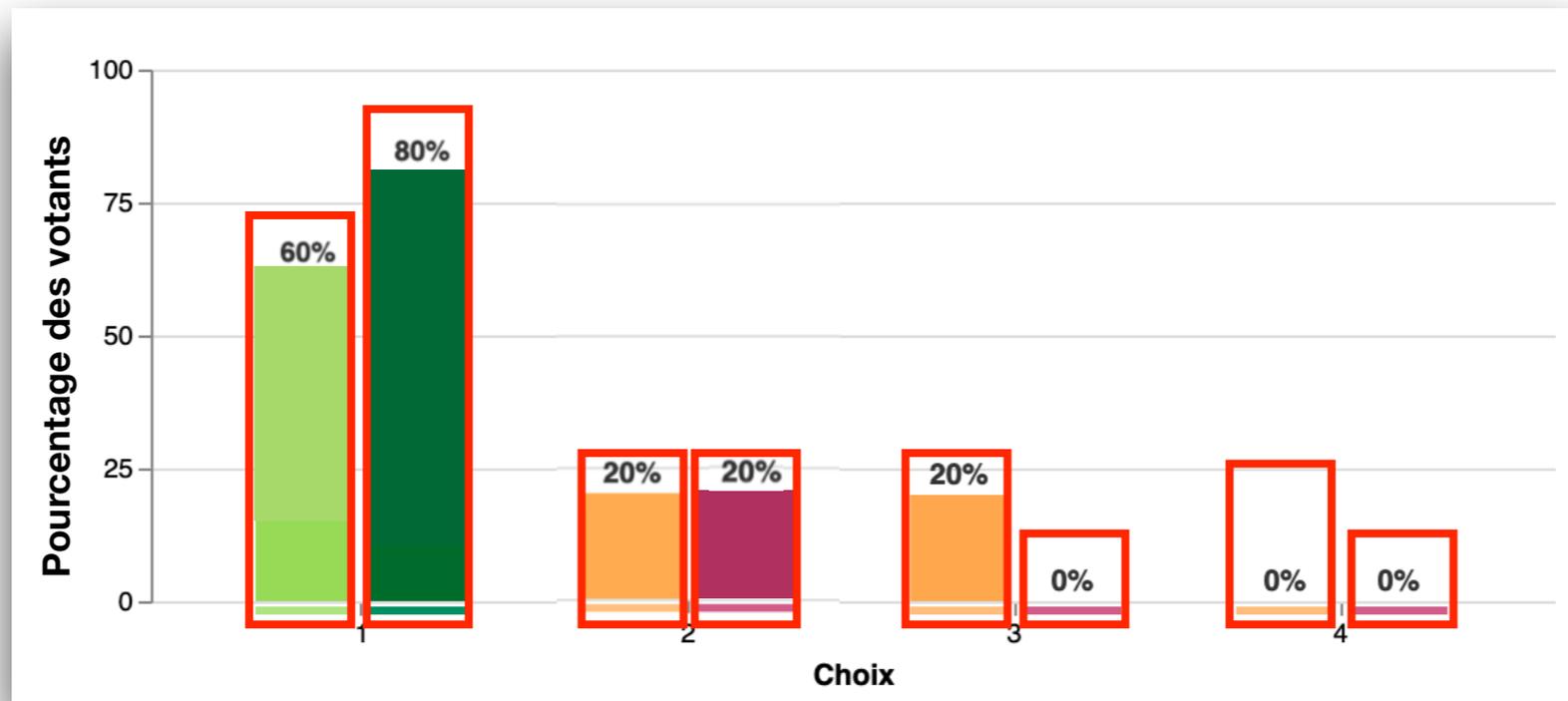
Les apprenants confrontent leurs points de vue



Les apprenants répondent à la question



Discussion



L'enseignant pose
une question



Les apprenants répondent
à la question



Les apprenants confrontent
leurs points de vue



Les apprenants répondent
à la question



Discussion



3.4/5 5 évaluations

Erik Erik (@erik) - Confiant(e)

Autre justification écrite pour la réponse correcte

3.2/5 5 évaluations

Albert Ein (@aein) - Confiant(e)

Une explication fausse mais pour une bonne réponse

2.8/5 5 évaluations

Thom Sil (@tsil) - Confiant(e)

Justification écrite pour la bonne réponse

Learning Analytics

LAK 2021, April 11-15 2021, Newport Beach, CA, USA

Andriamiseza et al.

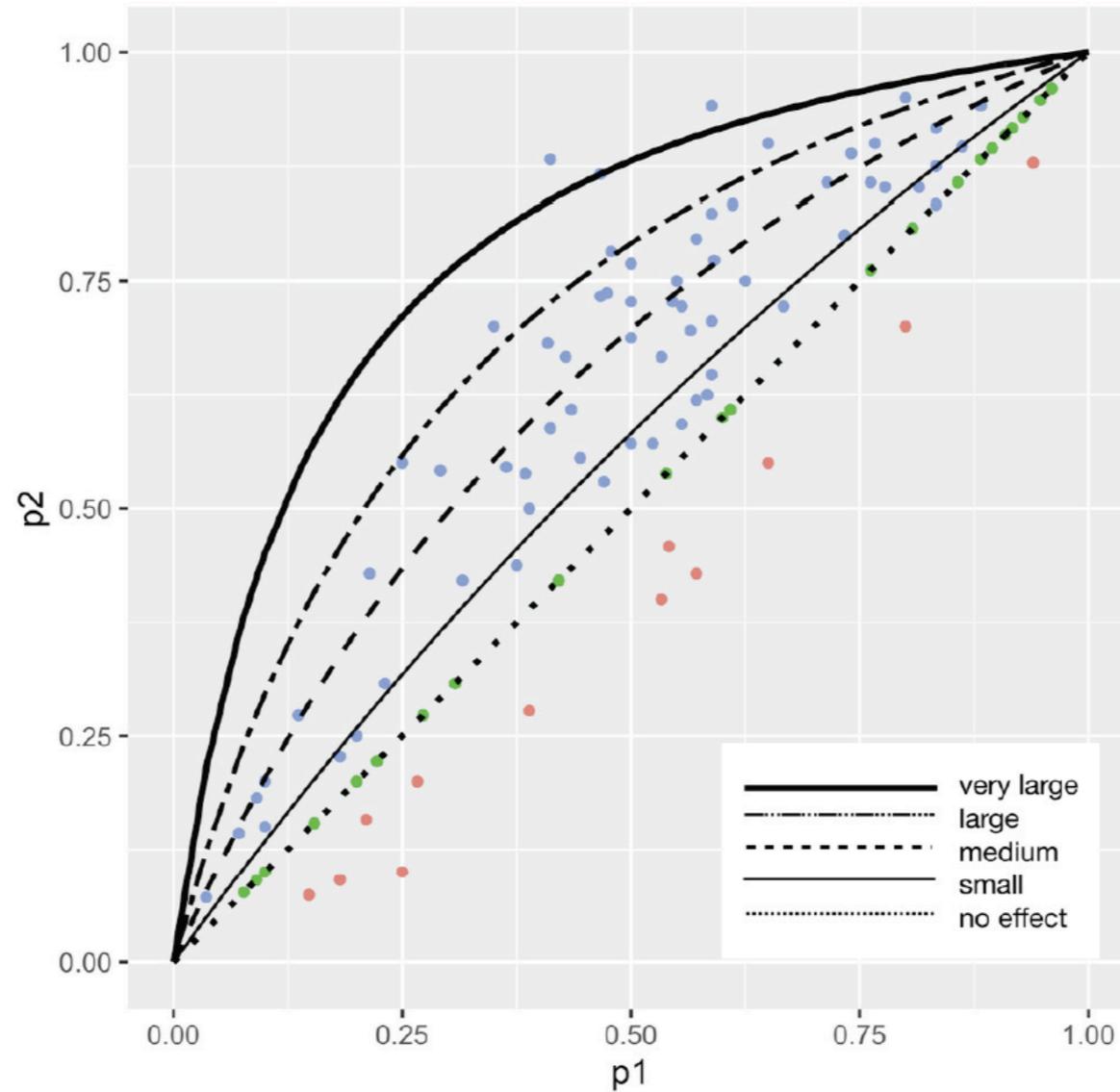


Fig. 3. Results of the peer discussion for the 104 sequences. Each point is the proportion of correct students' answers to a given question. Lines are isovalues of the Cohen's d effect size: no effect ($d = 0$), small ($d = 0.2$), medium ($d = 0.5$), large ($d = 0.8$), and very large ($d = 1.2$) effect. Red, green and blue points represent respectively harmful, inefficient and efficient sequences.

Bénéfices d'une évaluation par les pairs simpliste

Choix [4]
Si g(4) retourne 8
Votre évaluation :

Pas du tout d'accord

Je ne me prononce pas 1 2 3 4 5

- Pour l'évaluateur
 - Critère simple à comprendre
 - Interface de délibération très rapide : 1 click
 - Suffisant pour engager une posture réflexive sur les réponses
- Pour l'évalué et l'enseignant
 - Obtention d'un *feedback* agrégeant plusieurs jugements
- Pour l'enseignant
 - La phase d'évaluation par les pairs s'effectue dans une durée raisonnable

Limites d'une évaluation par les pairs simpliste

Choix [4]

Si g(4) retourne 8

Pas du tout d'accord

Votre évaluation :

Je ne me prononce pas

1

2

3

4

5

- Pour l'évaluateur
 - Ne guide pas l'évaluation
 - Très peu d'expressivité
 - Ne permet pas de justifier le "jugement"
- Pour l'évalué et l'enseignant
 - *Feedback* ramené à une valeur numérique
 - Pas de *feedback* textuel
- Pour l'évalué
 - Pas de guidage sur ce qui pourrait être fait pour améliorer la réponse

Questions de recherche

Comment améliorer l'évaluation par un pair d'une réponse écrite apportée à une question de compréhension ?

DRA XO

- Un cadre pour l'évaluation par les pairs d'une réponse écrite offrant
 - Une grille générique et dynamique de maximum 5 critères ordonnés
 - Des principes régissant le processus d'une évaluation
 - Un questionnaire dynamique type
 - Un calcul de score d'une réponse
 - Des garanties induites

DRAXO - Les 5 critères ordonnés

Une “bonne” réponse écrite doit être

- **D**échiffrable, compréhensible
- **R**épond à la question posée, n’est pas hors-sujet
- **A**ceptable, susciter l’adhésion, l’accord
- **eX**haustive, complète
- **O**ptimale

DRAXO - Les 2 principes du processus d'évaluation

Si un critère n'obtient pas une valeur suffisante, alors

- Les critères suivants ne sont pas évalués
- L'évaluateur doit justifier par écrit la valeur affectée à ce critère

Besoin d'un questionnaire dynamique !

Le questionnaire dynamique type

- Questionnaire Google Forms comme “démonstrateur”
 - <https://forms.gle/hanCSwVn4UW6uA6A6>
- Librairie Javascript pour intégration dans une application Web
 - En cours de développement
 - Première intégration : Elaastic prévue pour automne 2022

Calcul du score d'une réponse

- Soit Rep une réponse à évaluer

$$S(Rep) = [(V_D(rep) - 1) \times E_D(Rep) + (4 - V_R(Rep)) \times E_R(Rep) + (V_A(Rep) - 1) \times E_A(Rep) + (4 - V_X(Rep)) \times E_X(Rep) + (4 - V_O(Rep)) \times E_O(Rep)] \times \frac{10}{15}$$

Où

$$E_{\text{Criterium}}(Rep) = \begin{cases} = 0 & \text{si Criterium non évalué} \\ = 1 & \text{si Criterium évalué} \end{cases}$$

$$V_{\text{Criterium}}(Rep) = \begin{cases} = 1 & \text{si Criterium évalué à "Pas du tout d'accord"} \\ = 2 & \text{si Criterium évalué à "Pas d'accord"} \\ = 3 & \text{si Criterium évalué à "D'accord"} \\ = 4 & \text{si Criterium évalué à "Tout à fait d'accord"} \end{cases}$$

Calcul du score d'une réponse

- Exemples de score :
 - Réponse remplissant tous les critères évalués au mieux : $S(\text{Rep}) = 10$
 - Réponse tout à fait déchiffrable mais hors sujet : $S(\text{Rep}) = 2$
 - Réponse tout à fait acceptable mais pas du tout optimale : $S(\text{Rep}) = 8$

Les garanties induites

- Pour l'évaluateur
 - Un guide pour évaluer de manière plus précise une réponse écrite
 - La possibilité de fournir une explication sur **le** critère évalué comme défaillant
- Pour l'évalué et l'enseignant
 - Un *feedback* **textuel** sur le critère évalué comme défaillant par un évaluateur
- Pour l'enseignant
 - le temps reste maîtrisé
 - Sur une "bonne" réponse, un clique par critère : c'est rapide
 - Sur une "mauvaise" réponse : une seule explication ciblée sur un seul critère est attendue

Quelques hypothèses à tester

- Les évaluations effectuées dans le cadre DRAXO sont cohérentes d'un évaluateur à l'autre
 - Entre pairs
 - Entre pairs et enseignants
- Les *feedback* induits sont utiles à l'apprentissage
 - De l'évaluateur
 - De l'évalué
- Les *feedback* induits sont utiles à l'enseignant
 -



DRAXO

Un cadre pour l'évaluation formative par les pairs

Franck Silvestre - franck.silvestre@irit.fr

06/2022