

TEST EN ARGUMENTATION

YVES HÉBERT ET FLORIAN PÉLOQUIN, CÉGEP JOLIETTE-DE LANAUDIÈRE

Dans le cadre d'un projet subventionné par PAREA et le collège Joliette-De-Lanaudière, nous avons construit un test à choix multiples afin d'évaluer les compétences argumentatives des élèves. En regard de l'élaboration du test, nous présenterons quelques problèmes théoriques et leurs conséquences pratiques. De plus, nous indiquerons quelques avenues d'utilisation pédagogique du test.

Dans le cadre d'un projet, subventionné par PAREA et le Collège Joliette-De Lanaudière, nous avons construit un test à choix multiples afin d'évaluer les compétences argumentatives des élèves. Au regard de l'élaboration du test, nous en présenterons les principales étapes, quelques problèmes méthodologiques rencontrés et discuterons des questions théoriques que ceux-ci soulèvent, particulièrement avec l'évaluation de la suffisance.

Ce test, pour l'essentiel, se veut un instrument d'évaluation de la compétence argumentative. Il vise à identifier les forces et les faiblesses des élèves du collégial en argumentation afin de mieux cibler les interventions pédagogiques. Ce test s'ajoute aux autres tests qui évaluent l'argumentation. À notre connaissance, en français, deux tests sont actuellement disponibles pour évaluer l'argumentation: Paré et Questar. On peut aussi ajouter le test Cogit qui évalue plusieurs éléments de la compétence argumentative. Ces trois tests ont été développés dans le cadre de la recherche Apprentissage philosophique en réseau informatique de Robert Tremblay et Jean-Guy Lacroix.¹

Les compétences argumentatives telles qu'utilisées dans la vie de tous les jours ne se distinguent pas complètement des compétences langagières et des compétences logiques. Par exemple, le fait de reconnaître la conclusion

dans un texte argumentatif reste lié en partie à une compétence langagière. D'un autre côté, reconnaître une conclusion ou un argument sous-entendu est très souvent lié à une compétence liée au raisonnement logique. Il faut bien voir que lorsqu'on évalue les compétences argumentatives, on évalue en même temps des compétences langagières et logiques.

Une dimension d'emblée reconnue comme étant du domaine de la compétence argumentative est l'évaluation de l'acceptabilité des arguments et de la suffisance du lien entre les arguments et l'opinion défendue.

D'ailleurs, en philosophie, la formation collégiale insiste particulièrement sur cet aspect. Dans notre présentation, nous aborderons cet élément qui soulève des problèmes dans l'élaboration du test.

PRINCIPALES ÉTAPES D'ÉLABORATION DU TEST

Pour construire notre test, nous nous sommes référés aux manuels scolaires utilisés à travers le réseau, aux tests disponibles, au Québec et aux États-Unis, sur l'argumentation et sur la pensée critique.

Pour préciser le domaine et ses composantes, concevoir le test et rédiger les questions, nous avons rédigé un questionnaire afin de recueillir, auprès de professeurs intéressés à l'argumentation, leur conception de l'argumentation et les difficultés rencontrées par les élèves. Huit professeurs de cinq collèges ont répondu à l'invitation. Nous avons aussi tenu compte des commentaires de trois professeurs de notre département qui ont participé à la validation du questionnaire.

À la suite de l'analyse des réponses reçues, nous avons rédigé un tableau de spécification des composantes devant servir à une première version du test. Nous avons consulté trois professeurs d'université, experts en argumentation, afin de recueillir leurs commentaires.

La première version du test a été soumise à quatre professeurs de cégep, experts en argumentation, pour en valider le contenu et les choix de réponses. Les commentaires ont été recueillis soit à l'aide d'une fiche d'évaluation et de questions précises insérées dans le test,

¹. TREMBLAY, Robert, LACROIX, Jean-Guy, 1996, *Apprentissage philosophique en réseau informatique*. Montréal, Cégep du Vieux-Montréal, Service de la recherche, 231p.

soit par des rencontres avec deux d'entre eux. On peut noter que ces commentaires ont mis en relief la difficulté d'établir un choix de réponses multiples pour l'évaluation de la suffisance. Nous y reviendrons. Par ailleurs, nos conseillers en mesure et évaluation ont analysé, d'un point de vue docimologique, cette version du test.

La deuxième version a été soumise à un test de lisibilité auprès de soixante élèves du Cégep Joliette-De Lanaudière et auprès de six professeurs de notre département de philosophie.

La troisième version a été soumise à un professeur de français familier avec l'argumentation et a été validée par nos conseillers en mesure et évaluation. Cette version a été soumise à 437 élèves du Cégep Joliette-De Lanaudière.

Les analyses statistiques des résultats de l'expérimentation

indiquent que le test a certaines faiblesses au plan de la fidélité. Il a alors été décidé de réaliser un test-retest dans le but de vérifier la stabilité dans le temps des réponses des élèves. L'échantillon est composé de 112 élèves du Cégep Joliette-De Lanaudière.

Le test-retest confirme que le test n'a pas un très haut niveau de fidélité. Plusieurs hypothèses ont été soulevées. L'une d'entre elles a trait à la possibilité que la grille de réponses proposée ne soit pas tout à fait adéquate. Nous avons donc procédé de façon systématique à une nouvelle corroboration de la grille auprès d'un plus grand nombre d'enseignants et enseignants en philosophie.

C'est l'analyse des résultats de ce procédé de corroboration que nous présentons. Au regard de cette évaluation nous relèverons certains problèmes théoriques qui ont des conséquences pratiques.

TABLEAU 1: ÉTAPES MÉTHODOLOGIQUES

VERSIONS	VALIDATION
1 ^{RE} : 72 QUESTIONS; 4 SECTIONS	4 PROFS DE CÉGEP
2 ^{ME} : 70 QUESTIONS; 4 SECTIONS	TEST DE LISIBILITÉ (60 ÉLÈVES)
3 ^{ME} : 77 QUESTIONS; 5 SECTIONS	EXPÉRIMENTATION (437 ÉLÈVES)
4 ^{ME} : 57 QUESTIONS; 5 SECTIONS	TEST-RETEST (112 ÉLÈVES)
	CORROBORATION (28 PROFS)

ANALYSE DES RÉSULTATS DE LA CORROBORATION

L'objectif de ce procédé est de faire vérifier par des professeurs de différents cégeps, enseignant l'argumentation, la grille de correction utilisée.

Une lettre a été expédiée aux coordonnateurs des départements de philosophie leur demandant de transmettre le test aux professeurs qui donnent le premier cours de philosophie. De plus, un envoi individuel a été fait auprès d'enseignants en philosophie ayant un intérêt particulier pour l'argumentation. Nous demandions aux enseignants d'indiquer leur réponse à chacune des questions.

Au total, 29 professeurs de philosophie provenant de 18 cégeps ont participé à ce procédé de corroboration. Un des professeurs n'a pas répondu au test, mais nous a fait part de ses commentaires. Un autre professeur n'a pas répondu au sous-test II en indiquant que l'échelle était trop

subjective. Trois d'entre eux ont commenté certains aspects du test. À remarquer que seulement deux cégeps de l'île de Montréal sont ici représentés.

Le test est construit en tenant compte des cinq éléments de compétence suivants: reconnaissance de l'opinion défendue, évaluation de la suffisance, évaluation de l'acceptabilité, reconnaissance de l'opinion défendue sous-entendue et reconnaissance de l'argument sous-entendu. Ces éléments font chacun l'objet d'un sous-test.

Pour chacun des sous-tests, nous présentons un exemple des questions et les réponses recueillies pour chacune des questions. De plus, nous analysons les résultats. Nous considérons qu'un seuil de 25 bonnes réponses est satisfaisant pour valider une question.

SOUS-TEST I RECONNAISSANCE DE L'OPINION DÉFENDUE

EXEMPLE

Isabelle: «La plupart des festivals d'été au Québec sont déficitaires et il est inadmissible d'accepter des déficits pour de telles activités. Même si les Médiévales sont très populaires, il n'en demeure pas moins qu'elles ne font pas leurs frais. Par conséquent, on ne devrait pas financer, à même les fonds publics, la tenue de cette activité.»

- A) La plupart des festivals d'été au Québec sont déficitaires.
- B) Il est inadmissible d'accepter des déficits pour de telles activités.
- C) Même si les Médiévales sont très populaires, il n'en demeure pas moins qu'elles ne font pas leurs frais.
- D) Par conséquent, on ne devrait pas financer, à même les fonds publics, la tenue de cette activité.

Réponse

A B C (D) E

TABLEAU 2 RÉPARTITION DES RÉPONSES AU SOUS-TEST I SUR LA RECONNAISSANCE DE L'OPINION DÉFENDUE

I		L'opinion défendue par l'interlocuteur est:				
1	A	B				
	-	28				
2	A	B	C			
	-	-	28			
3	A	B	C			
	-	-	28			
4	A	B	C	D		
	28	-	-	-		
5	A	B	C	D	E	
	1	-	-	-	27	
6	A	B	C	D		
	2	-	-	26		
7	A	B	C	D		
	-	28	-	-		
8	A	B	C	D		
	-	27	-	1		
9	A	B	C	D		
	22	2	4	-		
10	A	B	C	D	E	
	-	28	-	-	-	
11	A	B	C	D	E	
	-	4	-	-	24	

Pour la reconnaissance de l'opinion défendue, nous avons retenu certains éléments de difficulté: la présence ou l'absence de marqueurs de relation, la place de l'opinion défendue et certains procédés «rhétoriques» (par exemple, un contexte de dialogue). Les questions ont été rédigées dans un français correct et selon une échelle de difficultés.

Pour 9 questions sur 11 (soit les questions 1, 2, 3, 4, 5, 6, 7, 8 et 10) au moins 27 répondants ont bien identifié la conclusion.

À la question 9, six personnes ont donné une autre réponse: soit 2 pour B et 4 pour C. Les quatre personnes qui ont répondu B au numéro 11 ne sont pas les mêmes qui ont répondu B ou C au numéro 9. Un répondant semble avoir des difficultés particulières: il s'est trompé à 6, 8, 9 et 11.

Nous reconnaissons à la question 9 sa difficulté particulière, ce pourquoi nous croyons qu'il est nécessaire de la conserver. Pour la question 11, nous pensons que le manque d'attention à la lecture de la consigne et une

ambiguïté dans la reprise de la consigne sur la feuille-réponse expliquent en partie que quatre personnes répondent B. En corrigeant la formulation sur la feuille-réponses, nous maintiendrons cette question.

Les résultats des divers procédés de validation accèdent la valeur du premier sous-test portant sur la reconnaissance de l'opinion défendue. Le sous-test donne des indices sur les capacités des élèves à bien lire un texte argumentatif. Comment faire une bonne évaluation de l'argumentation quand on ne peut reconnaître l'opinion

défendue? La recherche de la conclusion dans un texte argumentatif n'est pas aussi simple qu'on peut le croire. Les experts en argumentation arrivent de façon satisfaisante à trouver la conclusion, mais par contre certains enseignants non familiers avec l'enseignement de l'argumentation avaient quelques difficultés à repérer la conclusion de certaines questions. Naturellement, cette section pourrait être améliorée, mais elle reste, avec les corrections à apporter, très respectable et d'un niveau de difficulté acceptable.

SOUS-TEST II ÉVALUATION DE LA SUFFISANCE

DIRECTIVE:	Dans chacune des argumentations suivantes, en supposant que les arguments sont vrais, encerclez la lettre correspondant à la force des arguments.
P	des ARGUMENTS SANS AUCUN POIDS; les raisons énoncées ne justifient aucunement l'opinion défendue. Exemple: Nadia est la meilleure joueuse de tennis de la ville, car elle a des vêtements portant la marque Adidas.
F	de FAIBLES ARGUMENTS; les raisons énoncées justifient faiblement l'opinion défendue. Exemple: Nadia est la meilleure joueuse de tennis de la ville, car elle joue depuis son enfance.
B	de BONS ARGUMENTS; les raisons énoncées justifient modérément l'opinion défendue. Exemple: Nadia est la meilleure joueuse de tennis de la ville, car c'est l'avis de la plupart des amateurs.
T	de TR'S BONS ARGUMENTS; les raisons énoncées justifient fortement l'opinion défendue. Exemple: Nadia est la meilleure joueuse de tennis de la ville, car elle a gagné la majorité des compétitions des deux dernières années.

EXEMPLE	
Cette forme d'organisation du travail sera excellente pour notre entreprise, car elle est la plus récente qui nous a été proposée.	
Réponse	P (F) B T

NOTE: Les arguments sont imprimés en caractères gras et l'opinion défendue est imprimée en caractères normaux.

TABLEAU 3 RÉPARTITION DES RÉPONSES AU SOUS-TEST II SUR L'ÉVALUATION DE LA SUFFISANCE

II EN SUPPOSANT QUE LES ARGUMENTS SONT VRAIS, la force des arguments est: (P) ARGUMENTS SANS AUCUN POIDS, (F) FAIBLES ARGUMENTS, (B) BONS ARGUMENTS, (T) TR'S BONS ARGUMENTS.										
12	P	F	B	T		22	P	F	B	T
	-	9	10	8	1*		2	10	11	4 1
13	P	F	B	T		23	P	F	B	T
	2	21	4	-	1		3	13	10	1 1
14	P	F	B	T		24	P	F	B	T
	19	3	1	4	1		19	5	2	- 11?
15	P	F	B	T		25	P	F	B	T
	-	8	9	9	2		11	8	5	3 1
16	P	F	B	T		26	P	F	B	T
	4	15	7	1	1		6	13	8	- 1
17	P	F	B	T		27	P	F	B	T
	7	18	2	-	1		4	18	4	- 2
18	P	F	B	T		28	P	F	B	T
	19	6	2	-	1		14	11	2	- 1
19	P	F	B	T		29	P	F	B	T
	8	7	1	11	1		5	11	9	2 1
20	P	F	B	T		30	P	F	B	T
	2	7	11	7	1		21	2	-	3 2
21	P	F	B	T		31	P	F	B	T
	10	14	3	-	1		2	5	18	2 1

* Les chiffres de cette colonne désignent les non-répondants à la question.

La plupart des questions de ce sous-test illustrent des types de preuves couramment utilisés et que plusieurs considèrent comme des sophismes. Les autres font appel à des raisonnements logiques. Pour les besoins de l'évaluation de la suffisance, nous nous référons à des critères relevant de l'esprit critique. Pour les types de preuves couramment utilisés s'apparentant aux sophismes, contrairement à plusieurs qui les considèrent sans valeur, nous leur accordons un certain poids. Pour les questions rédigées selon les formes du syllogisme et dont la valeur de vérité de la conclusion serait indéterminée, on peut tout de même reconnaître aux prémisses un certain poids.

Selon notre grille de correction et en considérant un seuil de 25 bonnes réponses seule la question 13 peut être considérée valide.

Avant l'analyse des réponses, il importe de rappeler le

cheminement dans la détermination des réponses. Nous voulions, pour chacune des questions, qu'il n'y ait qu'une seule réponse possible. Lors de l'étape de l'expérimentation, nous avons quand même décidé de considérer pour certaines questions deux bonnes réponses, car nous n'arrivions pas à opter pour une seule réponse. Les analyses statistiques des résultats de l'expérimentation nous ont conduit en fin de compte à considérer pour toutes les questions deux choix de réponses. Les divers procédés d'élimination des questions ont fait en sorte que, à l'exception de la question 25, seules les questions ayant pour bonnes réponses (F) FAIBLES ARGUMENTS et B (BONS ARGUMENTS) ont été conservées.

En observant le tableau 3, on peut remarquer la diversité des réponses, par exemple la question 13 rejoint une grande majorité (21) des répondants pour une même réponse,

tandis que la question 19 illustre un étalement. Au seuil de 25, seule la question 13 peut être conservée.

Nous analysons huit questions qui illustrent la variété des réponses obtenues et certains problèmes théoriques rencontrés.

Tableau 4 Échantillon des réponses obtenues.

13. Il ne faut jamais se fier aux critiques de cinéma qui écrivent dans les journaux. **En effet, il m'est arrivé à quelques reprises de suivre les conseils de l'un d'entre eux à propos de supposés excellents films et de m'ennuyer à mourir pendant deux heures.**

P	F	B	T
2	21	4	-

17. **Malgré les faits accablants rapportés par un des témoins, je peux conclure, contrairement au journaliste de La Presse, que l'accusé est innocent. Tout le monde sait que ce témoin a des liens étroits avec le monde du crime organisé.**

P	F	B	T
7	18	2	-

18. **Les anthropologues ont constaté que les populations primitives vivant aux abords des lacs avaient un fort taux de natalité.** Cela confirme que la consommation de certains poissons avait un effet positif sur la fertilité des couples.

P	F	B	T
19	6	2	-

19. **Tous les produits de cette compagnie ont des éléments toxiques. Je sais que certains éléments toxiques ne sont pas illégaux.** Donc, certains produits de cette compagnie ne sont pas illégaux.

P	F	B	T
8	7	1	11

21. Il serait raisonnable de croire en l'existence d'êtres intelligents en dehors de notre système planétaire, **puisque la NASA a envoyé dans l'espace une sonde spatiale dans le but de capter des signaux que ceux-ci pourraient émettre.**

P	F	B	T
10	14	3	-

22. Cet été la mode sera aux couleurs pastels, je le sais, **vu que je travaille dans un magasin de vêtements.**

P	F	B	T
2	10	11	4

24. **Dans l'éventualité où le professeur est en retard, le règlement oblige les étudiants d'attendre en classe au**

moins dix minutes après le début des cours. Le règlement devrait être modifié de telle façon que, si un ou des élèves sont absents, le professeur doive attendre au moins dix minutes leur arrivée éventuelle.

P	F	B	T
19	5	2	-

28. On ne peut pas accorder de crédit au commentateur qui a déclaré que Bidon est le chef politique qui a remporté ce débat. **Pour tout dire, leurs épouses sont amies depuis leurs études collégiales.**

P	F	B	T
14	11	2	-

La question 13 est une argumentation s'appuyant sur une généralisation. La plupart des experts s'accordent pour dire que le nombre de conseils sur lesquels l'interlocuteur se base apporte un poids faible, mais n'est pas nécessairement suffisant pour en faire un bon ou très bon argument.

L'argument de la question 17 attaque la crédibilité du témoin. Ce qui nous semble étrange est le fait que 7 répondants n'accordent aucun poids à cette attaque; alors que c'est là un élément important pour évaluer une preuve en cour. À la question 28, un plus grand nombre de répondants, soit 14, n'accordent aucun poids à l'attaque contre la personne. Comment comprendre que, pour ces répondants, les liens d'amitié entre les épouses n'aient pu jeter un léger doute sur la crédibilité du commentateur? Peut-être qu'une telle révélation dans les journaux, au moment de l'exercice, aurait modifié le choix de ces répondants.

À la question 18, 19 répondants considèrent que la corrélation n'a aucune chance de correspondre à une relation de causalité. Nous concédons que les possibilités étaient très faibles, mais de là à ne voir aucune pertinence, cela nous surprend. Peut-être que, si la corrélation avait été établie en fonction d'une baisse du taux de natalité, on aurait accordé un certain poids; car certains répondants se seraient rappelé l'affaire Minamata. Les connaissances sur ces questions ont pu influencer les répondants.

Ce qui étonne à la question 19, c'est l'opposition entre deux groupes de répondants. La question est une argumentation ayant la forme d'un syllogisme catégorique. Ici, le raisonnement est non valide: le moyen terme est deux fois particulier. La conclusion qui découle de ce raisonnement reste indéterminée quant à sa valeur de vérité. On peut supposer que la non-validité du raisonnement a pu influencer une bonne partie des répondants qui ne voient aucun poids à l'argument. Mais devait-on aller à l'autre extrême et considérer que les arguments étaient très bons?

Peut-être qu'ici il était difficile d'imaginer qu'une compagnie puisse produire seulement des produits qui seraient illégaux. Alors les arguments n'ont-ils pas été évalués au regard de l'adhésion des répondants à cette conclusion?

La question 21 est une argumentation qui s'appuie sur l'autorité (avis d'expert). Comment expliquer que 10 répondants n'accordent aucun crédit à la NASA. Faut-il penser que la NASA envoie une sonde juste pour le plaisir sans croire qu'au moins il y a une possibilité que sa mission soit un jour remplie? Peut-être est-ce la crédibilité de la conclusion qui pose problème: si on ne croit aucunement à l'existence des extra-terrestres, il est fort probable qu'aucun argument n'aura de poids.

La question 22 est un autre argumentation qui se réfère à l'autorité, la grande majorité rejoint nos deux réponses et cela en deux groupes quasi égaux. La question que l'on peut se poser est de savoir s'il est possible de préciser les échelons de l'échelle de telle sorte que ces deux groupes puissent se rejoindre dans une réponse unique. En fonction de justifications semblables, des répondants ont pu opter soit pour «faible argument», soit pour «bon argument».

La question 24 est une argumentation qui fait appel à la réciprocité. Il est intéressant de constater que 19 répondants ne voient aucun poids à l'argument. Il peut être révélateur d'indiquer le pourcentage des répondants-élèves (lors de l'expérimentation) à cette même question: 32 % n'accordent aucun poids, 30 % considèrent que l'argument est faible, 27 % y voient un bon argument et 11 % statuent qu'il est un très bon argument. On peut noter que la perspective change selon que l'on est professeur ou élève. Il faut croire que le rôle social peut influencer notre jugement.

Il est clair que l'échelle d'évaluation repose sur des critères tels qu'il soit impossible d'établir un choix de réponses acceptables pour rejoindre l'assentiment de l'ensemble des professeurs, même si on accepte deux réponses contiguës comme étant les bonnes réponses. Un moyen de pallier le problème serait d'établir une grille de correction dichotomique, c'est-à-dire que les trois réponses contiguës ayant le plus grand nombre de répondants seraient considérées comme bonnes. En retenant cette hypothèse, alors 18 questions sur 20 pourraient être conservées au seuil de 25 bonnes réponses, soit les numéros 12, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31.

SOUS-TEST III ÉVALUATION DE L'ACCEPTABILITÉ

EXEMPLE

La position des groupes opposés à la consommation de drogues suppose l'abolition de l'utilisation des drogues à des fins médicales. Il ne faut donc pas faire de dons à aucun de ces groupes.

Réponse

V (F) ?

TABLEAU 4 RÉPARTITION DES RÉPONSES AU SOUS-TEST SUR L'ÉVALUATION DE L'ACCEPTABILITÉ

III EN VOUS RÉFÉRANT À VOS CONNAISSANCES,								
V Vous pouvez déterminer que l'argument est vrai.								
F Vous pouvez déterminer que l'argument est faux.								
? Vous ne pouvez déterminer si l'argument est vrai ou faux.								
32	V F ?	- 22 6	36	V F ?	21 3 4	40	V F ?	1 18 9
33	V F ?	1 19 8	37	V F ?	3 21 4	41	V F ?	- 22 6
34	V F ?	23 1 4	38	V F ?	12 3 13	42	V F ?	- 21 7
35	V F ?	2 19 7	39	V F ?	- 26 2	43	V F ?	16 2 10

Pour l'évaluation de l'acceptabilité, l'interprétation qui en est faite se réfère aux connaissances généralement reconnues par des personnes avisées et aux valeurs qui font l'objet d'un large consensus dans notre société. L'évaluation de l'acceptabilité a posé moins de problèmes que celle de la suffisance, car les possibilités de réponses n'étaient que de trois: l'indéterminé comprend en plus de la possibilité de la non-évaluation toutes les nuances d'acceptabilité.

Ce sous-test comprend 12 questions, mais pour faire l'évaluation nous en retenons neuf, trois (34, 36 et 38) sont des leurres afin de varier les réponses.

Ces neuf questions illustrent des types de preuves

couramment utilisées et que plusieurs considèrent comme des sophismes.

Pour ces neuf questions, nous considérons que la réponse acceptable était «vous pouvez déterminer que l'argument est faux». On peut prendre en considération que ceux qui répondent «vous ne pouvez déterminer si l'argument est vrai ou faux» (?), ont un point de vue plus critique que ceux qui ont répondu «vous pouvez déterminer que l'argument est vrai» à des questions qui manifestement contenaient des arguments erronés.

Si nous considérons l'indétermination comme une bonne réponse, alors toutes les questions ont 25 bonnes réponses et plus.

SOUS-TEST IV RECONNAISSANCE DE L'OPINION DÉFENDUE

EXEMPLE

Les étudiants en train de répondre à ce test ont été admis au niveau collégial. Or, aucun de ceux qui ont été admis au niveau collégial n'a eu d'emploi cet été. Donc,

- A) certains qui ont eu un emploi cet été sont en train de répondre à ce test.
- B) aucun étudiant en train de répondre à ce test n'a eu d'emploi cet été.
- C) certains étudiants en train de répondre à ce test n'ont pas eu d'emploi cet été.
- D) tous les étudiants en train de répondre à ce test ont eu un emploi cet été.
- E) certains qui sont en train de répondre à ce test ont eu un emploi cet été.

Réponse

A (B) C D E

TABLEAU 5 RÉPARTITION DES RÉPONSES AU SOUS-TEST SUR LA RECONNAISSANCE DE L'OPINION DÉFENDUE SOUS-ENTENDUE

IV	EN VOUS RÉFÉRANT AU CONTENU DU TEXTE, l'opinion défendue qui découle le plus directement de l'argument ou des arguments est:						A	B	C	D	E
44	A	B	C	D	E	48	A	B	C	D	E
	1	-	26	-	- 1*	-	-	1	1	26	
45	A	B	C	D	E	49	A	B	C	D	E
	1	25	-	1	1	4	24	-	-	-	
46	A	B	C	D	E	50	A	B	C	D	E
	7	2	4	-	13 2	-	3	1	1	23	
47	A	B	C	D	E						
	-	-	1	26	- 1						

* Les chiffres de cette colonne désignent les non-répondants à la question.

Dans le cas de la reconnaissance de l'opinion défendue sous-entendue, nous avons rédigé les questions en fonction

des formes du syllogisme catégorique et du syllogisme hypothétique.

Quatre questions, soit 44, 45, 47 et 48, ont 25 bonnes réponses et plus. Les questions 49 et 50 ont respectivement 23 et 24 bonnes réponses. Seulement sept personnes ont donné la bonne réponse à la question 46.

Ce sont des personnes différentes qui ont répondu A à la question 49 et qui ont répondu B ou C à la question 50. La question 46 pose des problèmes d'interprétation dus à

l'ambiguïté de sa formulation. Elle sera supprimée. Pour la question 50, ce qui a pu jouer c'est l'interprétation à donner au «si...alors»: elle est souvent interprétée dans un sens non formel comme étant un «si et seulement si». Les questions 47, 48, 49 et 50 présentent les diverses possibilités d'utilisation du syllogisme hypothétique. Pour cette raison, nous croyons important de les conserver.

SOUS-TEST V RECONNAISSANCE DE L'ARGUMENT SOUS-ENTENDUE

<p>Acheter des cadeaux est une tradition à Noël. Or,</p> <p>A) toutes les sociétés ont des traditions. B) depuis des centaines d'années à Noël, on achète des cadeaux. C) au Québec, on est comme les autres sociétés. D) la fête de Noël est une tradition. E) les traditions sont à respecter.</p> <p>Donc, il faut acheter des cadeaux à Noël</p>					
Réponse	A B C D (E)				

TABLEAU 6 RÉPARTITION DES RÉPONSES AU SOUS-TEST SUR LA RECONNAISSANCE DU OU DES ARGUMENTS SOUS-ENTENDUS

V EN VOUS RÉFÉRANT AU CONTENU DU TEXTE, L'ARGUMENT SOUS-ENTENDU QUI RELIE LE PLUS DIRECTEMENT LE OU LES ARGUMENTS À L'OPINION DÉFENDUE EST:												
51	A	B	C	D	E							
	25	1	-	2	-	55	A	B	C	D	E	
							-	-	5	-	22	1
52	A	B	C	D	E	56	A	B	C	D	E	
	-	1	26	-	1		-	-	-	27	1	
53	A	B	C	D	E	57	A	B	C	D	E	
	-	-	17	10	- 1*		25	-	-	-	3	
54	A	B	C	D	E							
	-	28	-	-	-							

* Les chiffres de cette colonne désignent les non-répondants à la question.

Dans le cas de la reconnaissance de l'argument sous-entendu, nous avons rédigé les questions en fonction des formes du syllogisme catégorique, du syllogisme hypothétique et du syllogisme disjonctif.

Les questions 51, 52, 54, 56 et 57 ont 25 bonnes réponses et plus.

La question 53 est à éliminer car elle est mal formulée.

Malgré cinq réponses jugées incorrectes, en tenant compte de sa complexité accrue du fait qu'elle fait appel à deux formes de syllogisme, nous croyons que la question 55 demeure valable.

CONCLUSION

La question de l'évaluation de l'argumentation est au cœur de la compétence argumentative. Or, c'est essentiellement cette question qui pose problème. Les résultats mitigés des

sous-tests II et III peuvent s'expliquer par les facteurs suivants:

-La rédaction de courtes questions a pu faire en sorte que, à cause de l'imprécision du contexte, celles-ci soient comprises différemment selon les répondants.

-L'absence de paramètres acceptables par tous pour évaluer une argumentation constitue un facteur déterminant pour expliquer la diversité des réponses.

-Les connaissances générales de chacun jouent dans l'évaluation.

-Les valeurs de chacun influencent l'évaluation.

-La crédibilité donnée à la conclusion a pu jouer sur l'évaluation soit de la suffisance, soit de l'acceptabilité des arguments.

-Les arguments relatifs aux questions ayant la forme d'un syllogisme non valide sont toujours considérés par certains comme ayant aucun poids, tandis que pour d'autres, dont nous, ils peuvent selon la formulation avoir un certain poids.

-La valeur donnée à une argumentation ayant la forme d'un «sophisme», tel un appel à un grand nombre de personnes, sera nulle pour certains alors que, pour d'autres, dont nous, une telle argumentation peut avoir une certaine validité.

Nous n'avons pas procédé pour les professeurs à une expérimentation test-retest. Cependant notre propre expérience montre que pour une même question notre évaluation peut varier et parfois grandement.

Pour les sous-tests II et III, malgré les consultations auprès d'experts et de confrères, il reste à poursuivre des consultations pour s'assurer des possibilités de dégager un consensus sur les réponses acceptables. Au regard des variations de nos réponses et compte tenu des divergences théoriques entre spécialistes en argumentation, il n'est pas évident que l'on puisse établir une échelle unique qui satisfasse toutes les tendances. Est-il envisageable, sans opter pour une échelle dichotomique, d'en arriver à la rédaction de questions dont l'évaluation obtient le consensus des experts et ne varie pas dans le temps ?