

L'économie expérimentale comme outil pédagogique : quelle efficacité ?

Nicolas Eber*

Dans cet article, nous présentons les résultats d'une expérience pédagogique originale, menée sur 118 étudiants de première année de l'Institut d'Études Politiques de Strasbourg. Une séance de travaux dirigés portant sur la « tragédie des ressources communes » a été menée, pour une partie des étudiants à partir d'un jeu pédagogique, pour les autres de manière classique. Nous comparons les résultats qu'obtiennent les étudiants des groupes « classiques » et « expérimentaux » lors de tests consacrés spécifiquement au thème des ressources communes. Nous n'observons aucune différence entre les deux groupes d'étudiants lors des tests passés en fin de séance. Par contre, les étudiants des groupes expérimentaux obtiennent des résultats nettement meilleurs que leurs homologues des groupes classiques lors d'évaluations réalisées trois mois plus tard. Nos résultats plaident ainsi en faveur de l'hypothèse selon laquelle l'expérimentation permet une meilleure rétention du savoir à long terme.

enseignement - jeux pédagogiques

Experimental economics as a pedagogical tool: which efficacy?

In this article, we report the results of a pedagogical experiment involving 118 first-year undergraduate students. In a course on the "tragedy of the commons", some students (in "experimental" sections) participated in a brief classroom game while others (in "control" sections) received a traditional lecture-oriented presentation. We compare student achievement in tests specifically devoted to common resources. We find no significant difference between students in the experimental sections and those in traditional lecture-oriented sections in the answers to a multiple-choice test implemented at the end of the session. However, we do find a very strong effect on the results of evaluations made three months later, with students of experimental sections being significantly more successful. These results seem to indicate that the main advantage of experiments is to improve the long-term retention of knowledge.

teaching - classroom experiments

Classification JEL : A22, C92

* LARGE, IEP, Université Robert Schuman (Strasbourg III), e-mail : nicolas.eber@urs.u-strasbg.fr.

L'auteur tient à remercier Stéphane Bertrand, Régis Blazy et Bertrand Koebel pour leur aide lors de la conception et la réalisation de ce projet, ainsi que les deux rapporteurs de la Revue dont les nombreuses remarques et suggestions ont permis une amélioration substantielle de l'article.

Bien que servant en priorité d'autres objectifs (tester la théorie, produire des connaissances nouvelles, aider à la décision publique et privée), l'économie expérimentale peut également être un moyen d'améliorer la pédagogie dans l'enseignement de l'économie (Eber et Willinger [2005]). Depuis quelques années, on assiste à une montée en puissance de l'utilisation d'expériences dans les cours d'économie¹. Eber [2003] fait le point sur l'utilisation de la démarche expérimentale dans l'enseignement de l'économie, en discutant notamment de la portée et des limites de la méthode et en faisant un inventaire complet des ressources disponibles.

Une question empirique fondamentale concerne l'efficacité réelle de la méthode en termes pédagogiques. Autrement dit, l'utilisation d'expériences dans les enseignements d'économie porte-t-elle ses fruits en termes de compréhension et d'apprentissage des concepts par les étudiants ? Un certain nombre d'études se sont intéressées à la question. Pour le moment, les résultats sont contrastés (cf. Eber [2003]). Certaines concluent que l'inclusion de jeux pédagogiques améliore les performances des étudiants aux tests de connaissance et de compréhension alors que d'autres ne trouvent aucun effet. En tout cas, aucune étude n'a pour l'instant montré d'effet négatif.

Les études empiriques menées jusqu'à présent évaluent généralement la compréhension des étudiants à l'aide de tests réalisés tout de suite après les cours ou dans le cadre des examens finaux, c'est-à-dire au terme d'un processus de révision. Or, un des arguments souvent avancés en faveur des jeux pédagogiques est qu'ils permettraient une meilleure rétention du savoir à long terme. Cet argument, tiré des sciences de l'éducation, repose sur l'idée que des élèves ayant expérimenté eux-mêmes un concept en font un apprentissage plus profond, plus ancré et donc plus « durable ».

Dans le présent article, nous testons cette hypothèse de l'apprentissage « durable » en menant une expérience pédagogique sur 118 étudiants de première année de l'Institut d'Études Politiques de Strasbourg. Une séance de travaux dirigés portant sur la « tragédie des ressources communes » a été menée, pour une partie des étudiants à partir d'une expérience (le jeu du bien public), pour les autres de manière classique (sans expérience). Nous comparons les résultats qu'obtiennent les étudiants des groupes « classiques » et « expérimentaux » lors de tests consacrés spécifiquement au thème des ressources communes. Nous n'observons aucune différence significative entre les deux groupes d'étudiants lors des deux QCM administrés en fin de séance. Pour ce qui concerne le premier QCM portant sur des questions techniques, les étudiants des groupes expérimentaux obtiennent une note moyenne de 12,7/16 pratiquement identique aux 12,6/16 de leurs homologues des groupes classiques. Pour le second QCM, comprenant deux questions de compréhension générale sur le concept de ressource commune, les pourcentages de réponses correctes s'élevaient à 38,57 (question 1) et 58,57 (question 2) pour les groupes expérimentaux contre 39,58 (question 1) et 64,58 (question 2) pour les groupes de contrôle, des différences de nouveau non significatives.

1. Bien qu'en retard, les enseignants français se laissent peu à peu convaincre...

Par contre, les étudiants des groupes expérimentaux obtiennent des résultats nettement meilleurs que leurs homologues des groupes classiques lors d'évaluations réalisées trois mois plus tard. Dans le cadre de ces évaluations, les étudiants devaient notamment répondre à deux questions ouvertes très générales sur les ressources communes : « Expliquez le plus précisément possible ce qu'on appelle tragédie des ressources communes » (question A) et « Donnez un exemple précis et détaillé de tragédie des ressources communes » (question B). Une note sur 4 a été donnée à chaque étudiant sur chacune des deux questions. Les étudiants des groupes expérimentaux obtiennent une note moyenne de 1,84 pour la question A et 1,51 pour la question B, contre 0,81 (question A) et 1,08 (question B) pour leurs homologues des groupes de contrôle. Un modèle économétrique simple confirme ce résultat : la régression de la note sur la variable indicatrice d'appartenance à un groupe expérimental et un certain nombre de variables permettant de contrôler l'impact des caractéristiques individuelles des étudiants (âge, sexe, série du bac, etc.) établit que, toutes choses égales par ailleurs, appartenir à un groupe expérimental améliore la note d'environ 1,04 point (sur 5) à la question A et de 0,45 point (sur 5) à la question B. Ainsi, une différence très nette et hautement significative en faveur des groupes expérimentaux apparaît-elle dans ces évaluations menées trois mois après le déroulement de la séance, un résultat qui plaide en faveur de l'hypothèse selon laquelle l'expérimentation permettrait une meilleure rétention du savoir à long terme.

Notre étude propose trois apports principaux à la littérature sur l'efficacité pédagogique de l'expérimentation dans les cours d'économie. En premier lieu, elle est, à notre connaissance, la première à prendre véritablement en considération l'horizon temporel de l'apprentissage, en distinguant compréhension immédiate et apprentissage dans la durée². En deuxième lieu, elle se distingue des autres études par un fort degré de contrôle dans le dispositif expérimental : groupes formés selon l'ordre alphabétique, élèves d'un même même promotion, instructeur identique pour tous les groupes, etc. En troisième et dernier lieu, il s'agit, toujours à notre connaissance, de la première étude de ce type sur une population d'étudiants français.

La suite de l'article est organisée de la manière suivante. La section 1 propose une synthèse des résultats empiriques concernant l'efficacité des jeux pédagogiques et présente les deux hypothèses que nous cherchons à tester. La section 2 décrit précisément notre méthode alors que les résultats sont présentés et commentés dans la section 3.

2. Gremmen et Potters [1997] comparent également les résultats d'étudiants ayant suivi un enseignement classique avec ceux d'étudiants ayant suivi un enseignement fondé sur une expérience en évaluant les connaissances dans la foulée de la séance puis, à l'aide d'une interrogation surprise, trois semaines plus tard. Cependant, le délai de trois semaines entre les deux évaluations est nettement insuffisant pour pouvoir véritablement évaluer l'apprentissage dans la durée. Sans surprise, d'ailleurs, les différences entre le groupe expérimental et le groupe classique observées par les auteurs trois semaines après la séance reflètent exactement celles obtenues immédiatement dans la foulée de la séance.

1. L'efficacité des jeux pédagogiques : revue de la littérature et hypothèses

L'effet des jeux pédagogiques sur la qualité de l'apprentissage des étudiants est une question empirique. Plusieurs études ont effectivement cherché à mesurer l'impact de l'utilisation de jeux pédagogiques sur les résultats des étudiants aux examens ainsi que sur leur satisfaction par rapport à l'enseignement³. Sur un plan méthodologique, le protocole est toujours du même type : il s'agit de comparer les résultats de deux groupes d'étudiants, l'un ayant suivi un enseignement « classique », l'autre un cours similaire, mais incluant un ou plusieurs jeu(x) pédagogique(s).

En dépit de chiffres anecdotiques assez édifiants circulant parmi les utilisateurs de jeux pédagogiques⁴, force est de constater que les résultats des études empiriques cherchant à évaluer de manière précise et rigoureuse l'impact de la méthode expérimentale sont contrastés. Certaines études (Frank [1997], Gremmen et Potters [1997], Bell et Delemester [1999], Emerson et Taylor [2004], Dickie [2006]) concluent qu'inclure des jeux pédagogiques améliore la compréhension des étudiants et, par conséquent, leur note moyenne aux examens finaux. D'autres (Cardell et al. [1996], Mullin et Sohan [1999], Yandell [1999]) ne trouvent aucun effet significatif, mais aucune ne conclut à un effet négatif. Le tableau 1 présente de manière synthétique les principales études récentes en en donnant les grandes caractéristiques et les résultats les plus importants.

Tableau 1. Études empiriques concernant l'effet de l'utilisation de jeux pédagogiques

Étude	Dispositif expérimental	Taille de l'échantillon	Effet sur les performances des étudiants	Remarques
Cardell et al. [1996]	- Plusieurs expériences de marché - Séances en salle informatique	- environ 1 800 étudiants : 43 % dans les groupes expérimentaux, 57 % dans les groupes de contrôle	- Pas d'effet significatif	- Répartition aléatoire des étudiants dans les groupes - Enseignants différents

3. Nous ne nous intéresserons ici qu'aux résultats concernant l'impact des expériences sur les résultats des étudiants. Voir Eber [2003] pour une synthèse des résultats concernant l'effet des expériences sur la satisfaction des étudiants.

4. Par exemple, Holt [1999, p. 609] rapporte qu'à l'Université d'Amsterdam, « le taux de succès dans le cours d'introduction à l'économie a été réduit de 50 % l'année suivant l'introduction d'exercices expérimentaux obligatoires ».

Étude	Dispositif expérimental	Taille de l'échantillon	Effet sur les performances des étudiants	Remarques
Frank [1997]	- Un seul jeu : le jeu de la ressource commune	- 416 étudiants : environ la moitié dans les groupes de contrôle	- Effet positif : sur l'une des questions posées lors d'un QCM, le taux de réponses correctes est de 38 % pour les étudiants des groupes de contrôle, de 46 % pour les étudiants ayant simplement regardé l'expérience, et de 63 % pour les étudiants (volontaires) ayant participé à l'expérience	- Niveaux de classes différents - 7 instructeurs différents - Fort biais d'auto-sélection des étudiants dans les groupes, seuls des étudiants volontaires participant à l'expérience
Gremmen et Potters [1997]	- Un seul jeu, portant sur la macroéconomie internationale	- 38 étudiants : 19 dans le groupe expérimental et 19 dans le groupe de contrôle	- Effet positif : note moyenne de 8,79/12 pour le groupe expérimental contre 7,42/12 pour le groupe de contrôle	- Très petit échantillon - Affectation des étudiants aux groupes par tirage au sort - Classes de niveau différent - Instructeurs différents
Yandell [1999]	- Comparaison de deux promotions successives, l'une (1997) ayant suivi un cours qui intégrait deux jeux pédagogiques, la suivante (1998) ayant reçu un cours qui en intégrait six	- 66 étudiants : 31 (promotion 1997) + 35 (promotion 1998)	- Pas de différence significative : la note moyenne de la promotion 1998 n'est que très légèrement supérieure à celle obtenue par la promotion précédente	- Promotions différentes - Même instructeur - Programme et examen final identiques

Étude	Dispositif expérimental	Taille de l'échantillon	Effet sur les performances des étudiants	Remarques
Emerson et Taylor [2004]	- 11 jeux pédagogiques sur un semestre de microéconomie	- 300 étudiants : 59 dans les groupes expérimentaux et 241 dans les groupes de contrôle	- Effet positif : amélioration de la note plus forte (d'environ 11 à 12%) dans les groupes expérimentaux que dans les groupes de contrôle	- Peu d'étudiants dans les groupes expérimentaux - Pas d'affectation aléatoire des étudiants dans les groupes - 7 enseignants différents
Dieckle [2006]	- 7 jeux pédagogiques sur un semestre de microéconomie	- 108 étudiants : 80 dans les groupes expérimentaux et 28 dans le groupe de contrôle	- Effet positif : en moyenne, entre l'évaluation d'avant-semestre et l'évaluation de fin de semestre, les étudiants des groupes expérimentaux améliorèrent leur note de 3,60 (en nombre de réponses correctes) contre une amélioration moyenne de 1,86 pour les étudiants du groupe de contrôle	- Affectation aléatoire des étudiants dans les groupes - Même instructeur pour les différents groupes - Plusieurs jeux

Deux remarques peuvent compléter utilement le tableau 1. En premier lieu, Brauer et Delemeester [2001, p. 234] notent que même si aucun effet n'est décelé sur le résultat final des étudiants lors des épreuves d'économie, le bien-être global des étudiants a pu quand même s'accroître significativement si l'utilisation de jeux pédagogiques leur a permis d'apprendre plus vite les cours d'économie et, par conséquent, d'avoir plus de temps à consacrer aux autres cours et à leurs loisirs.

En second lieu, l'effet sur le résultat final aux examens n'est qu'une mesure très partielle et imparfaite de l'impact pédagogique des expériences. Les chercheurs en sciences de l'éducation ont clairement montré que les étudiants apprennent mieux lorsque l'enseignement prend une forme expérimentale. La démarche pédagogique expérimentale conduit généralement à une implication des étudiants qui deviennent « actifs » et interagissent (en-

tre eux et avec l'enseignant) pendant le cours. Avec les jeux pédagogiques, l'étudiant devient donc « acteur » et non plus simple « spectateur » du cours. Il en résulte un mode d'apprentissage différent, les étudiants comprenant mieux les concepts théoriques lorsqu'ils les ont expérimentés eux-mêmes ; c'est la *théorie de l'apprentissage expérimental* (« *experiential learning* »), développée notamment par Kolb [1984]5. De plus, l'efficacité pédagogique de la démarche expérimentale serait particulièrement nette lorsqu'on envisage l'*apprentissage de long terme*, c'est-à-dire ce que l'étudiant retient du cours bien après les examens. D'une manière générale, il retiendrait mieux (plus longtemps) les concepts enseignés à partir d'expériences. Les expériences et leurs résultats semblent effectivement « marquer » les étudiants davantage que les cours magistraux classiques6, hypothèse que nous soulevons précisément tester dans la présente étude.

Nous avons mené une expérience pédagogique sur les étudiants de première année de l'Institut d'Études Politiques de Strasbourg durant l'année universitaire 2004/2005. L'objectif de cette expérience était de tester les deux hypothèses suivantes :

Hypothèse H1 : L'utilisation d'expériences dans l'enseignement de l'économie améliore la compréhension à court terme des étudiants.

Hypothèse H2 : L'utilisation d'expériences dans l'enseignement de l'économie améliore l'apprentissage à long terme des étudiants.

Notons que les deux hypothèses ne sont bien évidemment pas exclusives. Nous avons vu que la littérature donne des résultats contrastés concernant l'hypothèse H1. Par ailleurs, aucune étude n'a véritablement testé l'hypothèse H27. Or, nous pensons qu'il s'agit là d'un point crucial puisque, du point de vue des sciences de l'éducation, l'argument fort en faveur des expériences est précisément qu'elles permettraient une rétention plus profonde, plus durable du savoir, en bref qu'elles marquent davantage les étudiants que les méthodes classiques.

2. Méthode

Les étudiants de première année de l'Institut d'Études Politiques de Strasbourg suivent tous un cours de microéconomie au premier semestre. Dans le cadre de ce cours obligatoire sont organisés, en parallèle, des travaux dirigés (appelés « conférences de méthode ») en petits groupes, d'une durée d'une heure et demi par séance. Les groupes sont définis en début d'année

5. Et au sujet de laquelle on se réfère parfois à la célèbre maxime de Confucius : « Dis-moi et j'oublierai, montre-moi et je me souviendrai, implique-moi et je comprendrai »...

6. A cet égard, on peut, par exemple, citer Frank [1997, p. 768] : « That a teaching method is appealing to students, as classroom experiments are, is of special importance for storage in the long-term memory [...] If this is true, students who merely followed the lecture would not only learn less, but also more than those who also watched the experiment ».

7. Comme nous l'avons déjà mentionné, nous pensons que l'expérience de Gremmen et Potters [1997], avec des évaluations dans la foulée de la séance puis trois semaines plus tard, ne permet pas de distinguer de manière satisfaisante les effets de court terme et de long terme dans l'apprentissage des étudiants.

selon l'ordre alphabétique des étudiants⁸. Notre expérience concerne 5 groupes d'une vingtaine d'étudiants, soit en tout 118 étudiants⁹ (sur les 150 que compte la promotion)¹⁰.

Durant le semestre, l'ensemble des cinq groupes a suivi exactement les mêmes séances de travaux dirigés de microéconomie, à une exception près. Lors de la dernière séance (dernière semaine de janvier 2005), consacrée au thème des ressources communes, trois groupes ont assisté à une présentation partant d'une expérience simple, alors que les deux autres groupes ont suivi une présentation standard (sans expérience), au contenu par ailleurs rigoureusement identique. En effet, l'instructeur (Stéphane Bertrand, alors ATER à l'IEP de Strasbourg) avait préparé, par écrit, les deux types de séances et nous nous étions assurés du plus grand parallélisme dans leur contenu.

Comme Frank (1997), nous avons opté pour le thème des ressources communes car nous savions qu'il n'avait pas été directement évoqué en cours magistral et que, par conséquent, le savoir des étudiants en la matière ne découlerait que du contenu de la séance de travaux dirigés qui lui était consacrée. Par ailleurs, restreindre l'expérience pédagogique à une seule séance (et donc un seul thème) nous paraissait souhaitable par rapport à une expérience se déroulant sur l'ensemble des séances (et qui aurait consisté à proposer systématiquement des jeux pédagogiques à trois groupes et des séances classiques aux deux autres). En effet, cette restriction à une seule séance nous permet de contrôler au mieux les effets extérieurs qui peuvent intervenir dans l'apprentissage des étudiants sur un semestre complet. Seul un dispositif avec séance unique, sur un thème non traité en cours magistral, nous permettait d'obtenir un contrôle satisfaisant sur l'ensemble des paramètres pouvant interférer avec l'expérimentation dans le mécanisme d'apprentissage des étudiants.

2.1. Les deux versions de la séance

2.1.1. La séance avec expérience

Dans la version dite « expérimentale », la séance consistait à démarrer d'une expérience classique : le jeu de contribution volontaire à un bien pu-

8. L'ordre alphabétique est le seul critère qui intervient. En effet, la 1^{re} année ne comporte aucun choix d'option et aucun horaire antéposé pour des sous-catégories d'étudiants (pas de créneaux réservés pour le sport, etc.). À quatre exceptions près, les étudiants ont respecté l'affectation alphabétique des groupes.

9. Cette population d'étudiants est composée à 56% de filles et a un âge moyen de 19,4 ans (écart-type : 0,75). Par ailleurs, la moitié des étudiants (50%) ont obtenu un bac scientifique (S), alors que la proportion de bacheliers ES et L sont, respectivement, de 33% et de 17%. La note moyenne obtenue au bac est de 14,44/20 (écart-type : 1,95). Enfin, notons qu'en raison de la spécificité de l'organisation des études à l'IEP de Strasbourg, qui recrute en 1^{re} année sur concours et qui, à l'époque, n'autorisait le redoublement qu'à titre exceptionnel, l'échantillon ne comporte que 5 redoublants.

10. Deux groupes de travaux dirigés n'ont pas été concernés par l'expérience dans la mesure où ils étaient pilotés par un autre enseignant.

blic. Rappelons-en rapidement la structure. Le jeu concerne quatre joueurs ne pouvant pas communiquer. Chacun des quatre joueurs reçoit 10 jetons et peut soit les conserver pour lui, soit contribuer à une cagnotte commune. Le gain de chaque joueur est égal à :

(2 euros x nombre de jetons conservés) + (1 euro x nombre total de jetons placés par tous les joueurs dans la cagnotte commune),¹¹

Ce « jeu du bien public » met en scène, de manière particulièrement simple, le problème du financement des biens collectifs, le phénomène d'externalité ou encore le comportement de passerager clandestin. Il s'agit d'un grand classique de l'économie expérimentale (cf. Eber et Willinger [2005, p. 88-90])¹². Le lien entre ce jeu et la tragédie des ressources communes est évident puisqu'il fait intervenir les principaux mécanismes qui la caractérisent¹³.

Les étudiants des groupes expérimentaux ont tous joué plusieurs répétitions du jeu du bien public, dans une version spécifiquement adaptée à l'optique pédagogique¹⁴. Le comportement des étudiants pendant le jeu s'est avéré conforme aux observations habituelles, avec des contributions moyennes comprises entre 30 et 50% de la dotation initiale selon les groupes et décroissantes au fil des répétitions du jeu.

Le déroulement de l'expérience s'est généralement étalé sur une vingtaine de minutes. À la fin de l'expérience, l'instructeur a fait le point sur les résultats, expliqué le jeu et démarré alors la présentation des concepts, en donnant les définitions d'une ressource commune et d'une externalité, en donnant des exemples, etc. À la fin de la séance, tous les étudiants ont dû répondre à deux QCM portant exclusivement sur le thème des ressources communes. Le premier QCM concernait spécifiquement le jeu du bien public alors que le second contenait deux questions plus générales sur les concepts théoriques présentés au cours de la séance.

On notera qu'aucun des étudiants concernés par l'étude (ni dans les groupes expérimentaux, ni dans les groupes classiques) n'avait participé à une expérience « réelle » (ou pédagogique) de bien public avant la séance et aucun n'a participé à une autre expérience entre la séance et l'évaluation de long terme trois mois plus tard.

11. Par exemple, si vous participez à ce jeu, que vous décidez de conserver 5 jetons et que les trois autres participants placent 12 jetons en tout, votre gain sera égal à : $(2 \times 5) + (1 \times 17) = 27$ euros.

12. Alors que la théorie standard prédit un comportement de passerager clandestin aboutissant à l'absence totale de contribution à la cagnotte, les observations expérimentales montrent, au contraire, des niveaux de contribution moyens compris entre 40 et 60% de la dotation initiale des sujets.

13. D'ailleurs, une variante du jeu du bien public est utilisée en économie expérimentale sous le nom de « jeu de la ressource commune », notamment dans le cadre de problématiques spécifiquement environnementales (cf. Eber et Willinger [2005, p. 90-92]).

14. Cf. Holt et Lury [1997] et Eber [2006] pour une présentation de cette version « pédagogique » du jeu du bien public.

2.1.2. La séance « classique » (sans expérience)

La séance dite « classique » suivait un déroulement similaire à la précédente, la seule différence résidant dans l'absence de réalisation de l'expérience. Afin d'assurer une comparabilité maximale des deux types de séance, la séance classique était calquée sur la séance expérimentale. En début de cours, l'instructeur (qui était bien entendu toujours le même, à savoir Stéphane Bertrand) expliquait le jeu du bien public mis en œuvre dans la séance expérimentale. La différence est qu'il ne faisait pas jouer les étudiants, mais se contentait d'une présentation et d'une explication *théorique* du jeu. Ensuite, le déroulement de la séance était rigoureusement identique, à savoir que les concepts, les exemples, etc., étaient présentés et discutés exactement de la même manière. À la fin de la séance, les étudiants se sont vus confrontés aux deux mêmes QCM que leurs camarades des groupes « expérimentaux », à savoir le QCM sur le jeu (dont ils avaient eu une présentation théorique, mais qu'ils n'avaient pas expérimenté) et le QCM sur les concepts.

La séance « classique » se contentait d'une présentation théorique du jeu, elle était légèrement plus courte que la séance « expérimentale ». Cependant, l'écart n'était pas très important (de l'ordre de 5 à 10 minutes) puisque l'expérience va très vite, son déroulement ne prenant qu'une vingtaine de minutes tout compris. Cela dit, nous avons convenu de maintenir exactement le même contenu dans les deux séances, quitte à libérer un peu plus tôt les étudiants des groupes classiques. Il s'agit là d'un point important. Nous avons testé le bénéfice potentiel en termes pédagogiques de la méthode expérimentale, sans chercher à en évaluer les coûts (en termes de temps passé en classe à faire l'expérience, de préparation pour l'instructeur, de difficultés d'organisation, etc.). Dans cette optique, l'approche est donc partielle. Dans des recherches futures, il conviendra d'évaluer l'arbitrage entre les bénéfices et les coûts de l'approche expérimentale.

2.1.3. Contrôle et comparabilité

L'une des grandes faiblesses de la littérature sur l'efficacité pédagogique des expériences est le manque de contrôle des chercheurs dans leurs dispositifs. Bien souvent, par exemple, la formation des groupes d'étudiants ne provient pas d'un tirage au sort ou de l'ordre alphabétique, mais résulte d'inscriptions volontaires des étudiants dans tel ou tel groupe. De plus, les différents groupes ont généralement des instructeurs différents. Par exemple, dans l'étude de Frank [1997], l'auteur agrége les résultats sur diverses classes correspondant à des universités différentes, à des niveaux d'étude différents et à des instructeurs différents¹⁵.

Une grande force de notre recherche réside précisément dans le fort degré de contrôle que nous avons pu instaurer tout au long du dispositif. Ce

contrôle et, par suite, la comparabilité entre les groupes expérimentaux et classiques, sont particulièrement forts puisque :

- les étudiants ont été répartis dans les groupes en fonction de l'ordre alphabétique et le choix des trois groupes expérimentaux s'est fait au hasard ;
- l'instructeur était le même pour tous les groupes (ce qui assure l'absence de tout effet d'instructeur) ;
- le thème de la séance (la tragédie des ressources communes) n'avait pas été vu en cours magistral ;
- le déroulement et le contenu des séances étaient rigoureusement identiques, la seule différence étant que le jeu du bien public était dans un cas joué (groupes expérimentaux), dans l'autre cas uniquement présenté de manière théorique (groupes classiques).

2.2. Les évaluations

2.2.1. L'évaluation de court terme

L'hypothèse H1 concerne l'impact de l'expérimentation sur la *compréhension immédiate* des étudiants. Pour mesurer cette compréhension immédiate, le format QCM nous paraissait approprié.

À la fin de la séance, les étudiants ont dû remplir deux QCM. Le premier contenait 16 questions portant spécifiquement sur le jeu du bien public et devait permettre de mesurer le degré de compréhension « technique » du jeu proprement dit. Le second QCM proposait deux questions de compréhension générale sur le thème de la séance, à savoir la « tragédie des ressources communes ». À des fins de comparabilité avec l'étude de Frank [1997], nous avons repris deux questions testées par cet auteur en nous contentant de les traduire¹⁶. La première question (Q1) était formulée de la manière suivante :

Q1. Parmi les cas suivants, lesquels relèvent d'une « tragédie des ressources communes » ?

- a) la pêche excessive dans les eaux internationales ;
- b) les embouteillages à l'heure de pointe ;
- c) deux collègues partagent un bureau, l'un fume et incommode l'autre ;
- d) un employé utilise avec négligence sa voiture de fonction, dont il est le seul utilisateur ;
- e) la déforestation dans certaines parties de l'Afrique pour produire du bois de chauffage ;

15. De plus, seuls des étudiants volontaires participent effectivement à l'expérience.

16. Cf. Frank [1997, p. 765-766] pour la version anglaise des questions.

f) la surpopulation dans le Tiers-Monde.

La question Q1 renvoie à des exemples de tragédie des ressources communes. Pour cette question, les bonnes réponses sont *a*, *b*, *c* et *f*¹⁷.

La seconde question était formulée comme suit :

Q2 : Imaginez un pâturage public dans un village de bergers. La viande, la laine et le lait sont vendus à des prix constants sur le marché mondial. Quelles mesures pourraient réduire à long terme la surexploitation du pâturage ?

- a) le nombre de bergers ayant le droit d'utiliser le pâturage est réduit de 50 à 30 ;
- b) le pâturage est agrandi de 50 % grâce à l'assèchement d'un marécage ;
- c) on explique aux villageois pourquoi il y a trop de moutons dans le pâturage ;
- d) le pâturage est morcelé afin que chaque berger en possède une petite parcelle (côturee).

La question Q2 concerne les méthodes permettant de « résoudre » la tragédie des ressources communes. La seule réponse correcte est la réponse *d*.

L'un des problèmes méthodologiques majeurs auxquels nous avons été confrontés est lié au fait que l'ensemble des étudiants n'avaient pas cours au même moment puisque les groupes étaient, bien entendu, espacés dans le temps (sur deux journées, le lundi et le mardi). Ainsi, il eût été possible que les étudiants des premiers groupes informent ceux des groupes suivants de la teneur de la séance et, surtout, de l'évaluation, faussant ainsi les résultats. Pour traiter de ce point délicat, nous n'avons pas trouvé d'autres solutions que de mentir aux étudiants¹⁸. L'instructeur leur a fait croire, au moment où ils répondaient aux QCM, que la note qu'ils obtiendraient serait prise en compte dans la moyenne (pour les inciter à répondre du mieux possible). Après le ramassage des questionnaires, il leur annonça qu'en fait il n'en serait rien, mais que l'évaluation lui serait simplement utile dans le cadre d'un travail de recherche sur l'enseignement de l'économie. A ce moment-là, il demanda aux étudiants la plus grande discrétion par rapport au contenu de la séance et des tests d'évaluation, « par respect pour le projet scientifique en cours »¹⁹. Le même dispositif a été utilisé pour tous les groupes. Nous pensons qu'il a été efficace puisque : (i) au moment de répondre aux QCM, les étudiants étaient persuadés de faire face à une interrogation surprise dont la note serait intégrée à leur moyenne, ce qui les incitait à répondre du mieux possible ; (ii) sachant, en sortie de séance, qu'en fait, la note ne serait pas prise en compte dans la moyenne et qu'il n'y avait donc pas d'enjeu majeur pour leurs camarades des groupes suivants et, de plus, probablement sensibles à la demande de discrétion « pour la

17. Comme le note Frank (1997, p. 766), la réponse *f* concernant la surpopulation prête à discussion.

18. Ce qui, soit dit en passant, est tout à fait contraire aux standards de la recherche en économie expérimentale (Eber et Willinger (2005, p. 23-24)).

19. Notons d'ailleurs que les étudiants s'orientent de la séance sans avoir les questions, ni les réponses aux différentes questions.

science » émanant de leur professeur, il est peu probable que les étudiants des premiers groupes aient cherché à « aider » ceux des groupes suivants²⁰.

2.2.2. L'évaluation de long terme

Aiors que les deux QCM administrés dans la foule de la séance devaient nous permettre de tester l'hypothèse H1 concernant la compréhension immédiate des étudiants, une autre évaluation, plus tard dans l'année, devait, elle, nous permettre d'envisager l'apprentissage de long terme et donc d'évaluer la pertinence de l'hypothèse H2.

Pour cela, notre protocole nous paraît tout à fait adéquat puisque les étudiants ne faisaient plus de microéconomie au second semestre, mais suivaient un enseignement d'introduction à la macroéconomie dans lequel aucune référence n'a été faite à la notion de ressources communes, ni dans le cours magistral (assuré par N. Eber), ni dans les travaux dirigés (toujours assurés par S. Bertrand). Les groupes de travaux dirigés n'ayant pas été modifiés entre le premier et le second semestre, nous avons pu, dans le cadre de l'avant-dernière séance de l'année (dernière semaine d'avril 2005), administrer un nouveau test, toujours exclusivement consacré au thème des ressources communes. Ce test se déroulait donc environ trois mois après la séance. Les étudiants n'ayant pas d'examen de microéconomie au second semestre, on peut légitimement supposer qu'aucun n'avait révisé le thème des ressources communes depuis trois mois.

Pour cette seconde évaluation, le QCM nous paraissait moins approprié. En effet, nous cherchions à mesurer la capacité des étudiants à restituer les points principaux (définition des concepts, exemples classiques) plutôt qu'à répondre à des questions spécifiques, la formule QCM permettant en outre aux étudiants de se remémorer les concepts au fur et à mesure qu'ils avancent dans les questions et qu'ils découvrent les différentes propositions. Pour neutraliser cet « effet de remémoration » en cours d'évaluation, nous avons opté pour un test de connaissance générale, consistant en deux questions ouvertes, à savoir :

QA : Expliquez le plus précisément possible ce qu'on appelle « tragédie des ressources communes ».

QB : Donnez un exemple précis et détaillé de « tragédie des ressources communes ».

Tous les étudiants des cinq groupes ont dû répondre à ces deux questions. Le test a été présenté comme une interrogation surprise, les étudiants devant indiquer leur nom sur la copie.

C'est l'auteur qui a corrigé les copies, en donnant une note comprise entre 0 et 4 pour chacune des deux questions. La notation s'est effectuée « en

20. Cette conjecture semble d'ailleurs vérifiée par l'absence de différences significatives dans les évaluations entre les premiers groupes et les derniers.

aveugle »²¹. Le correcteur connaissait le nom de l'étudiant (puisque celui-ci était indiqué sur la copie), mais pas son groupe de travaux dirigés²². Autrement dit, le correcteur ne pouvait pas savoir, au moment de la notation, si l'étudiant dont il corrigait la copie faisait partie d'un groupe expérimental ou d'un groupe classique. Il ne l'a découvert qu'après la correction de l'ensemble des copies.

À des fins de comparabilité avec la première évaluation, nous avons également, à la suite du test de connaissance générale, administré le même QCM que l'un de ceux proposés dans la foulée de la séance, à savoir celui portant sur les questions Q1 et Q2²³. Le QCM a été administré après avoir ramassé les copies sur les questions QA et QB car il était important que les étudiants répondent à ces deux questions sans aucun indice.

Concernant l'incitation des étudiants, nous avons reconduit le même dispositif que pour la première évaluation, même si la crédibilité de l'annonce d'une prise en compte de la note dans la moyenne, au moment de l'évaluation, était probablement dépréciée par ce qui s'était passé trois mois auparavant ! Cela dit, l'étudiant rendait une copie nominative à l'instructeur qui devait, quelques jours plus tard, lui donner une note de contrôle continu, et il avait donc tout intérêt à faire du mieux possible. Au moment où les questionnaires ont été distribués, nous avons fait croire aux étudiants que la note « pourrait être prise en compte dans la moyenne générale ». Comme pour la première évaluation et afin d'éviter une communication entre les étudiants des différents groupes, nous avons annoncé, après le ramassage des copies, qu'en fait l'évaluation avait de nouveau été organisée dans le cadre d'un travail de recherche sur l'enseignement de l'économie et que les notes ne seraient pas prises en compte dans la moyenne générale. Comme trois mois auparavant, nous avons demandé aux étudiants des premiers groupes la plus grande discrétion au sujet de l'évaluation « pour préserver la valeur scientifique du travail »²⁴. Des discussions avec les étudiants des derniers groupes nous laissent à penser que cette discrétion a effectivement été respectée par les étudiants des premiers groupes. D'ailleurs, aucune différence significative n'a pu être détectée entre les premiers et les derniers groupes, ni parmi les groupes classiques, ni parmi les groupes expérimentaux, que ce soit pour les réponses aux deux questions (QA et QB) du test de connaissance générale ou pour les réponses aux deux questions (Q1 et Q2) du QCM.

21. Bien entendu, cette note n'est pas sans révéler une certaine subjectivité. Cela dit, nous pensons que, sur l'ensemble de l'échantillon, le biais est faible.

22. Le correcteur n'assurant que le cours magistral et pas les travaux dirigés, il ne connaît aucun des étudiants nominativement.

23. Notons que les étudiants n'avaient pas eu le corrigé de ce QCM.

24. Comme trois mois plus tôt, les étudiants sortaient de la séance sans les formulaires d'évaluation et sans avoir eu la correction du test.

3. Résultats

Nous présentons d'abord les résultats des deux QCM administrés en fin de séance (3.1) puis ceux des évaluations effectuées trois mois plus tard (3.2).

3.1. Résultats des évaluations réalisées en fin de séance (hypothèse H1)

L'analyse des résultats des deux QCM administrés en fin de séance permet de tester l'hypothèse H1 sur l'impact pédagogique de *court terme* de l'expérimentation. Pour chacun des deux QCM, nous comparons les résultats des étudiants appartenant aux trois groupes expérimentaux avec ceux des étudiants faisant partie des deux groupes classiques²⁵.

3.1.1. Le QCM portant sur le jeu

Le QCM portant spécifiquement sur le jeu du bien public comprenait 16 questions. Le tableau 2 compare les résultats des groupes « expérimentaux » et des groupes de contrôle (« classiques »).

Tableau 2. Notes moyennes au QCM portant sur le jeu

Groupes expérimentaux		Groupes de contrôle	
Groupe	Moyenne de bonnes réponses (/16)	Groupe	Moyenne de bonnes réponses (/16)
Groupe 1	13,3	Groupe 4	12,6
Groupe 2	12,5	Groupe 5	12,6
Groupe 3	12,4		
TOTAL	12,7	TOTAL	12,6

Au total, les 70 étudiants des groupes expérimentaux ont obtenu une note moyenne de 12,7/16 (écart-type = 2,7/41) alors que les 48 étudiants des groupes

25. Nous avons choisi de présenter séparément, plutôt que de manière groupée, les résultats des deux QCM afin d'isoler, dans une optique de comparabilité, le second QCM. En effet, seul le second QCM a été repris dans l'évaluation trois mois plus tard. De plus, le second QCM comprend deux questions utilisées par Frank [1997] et permet donc une comparaison directe avec cette étude.

pes classiques ont obtenu la note moyenne de 12,6/16 (écart-type = 3,292), soit une différence non significative, quel que soit le test statistique utilisé²⁶.

3.1.2. Le OCM portant sur les concepts

Dans l'optique de répliquer l'étude de Frank [1997], nous avons calculé, comme lui, le pourcentage de réponses correctes, pour les groupes expérimentaux et les groupes de contrôle, à chacune des questions Q1 et Q2. Comme lui, nous comptons comme correctes les réponses ne comportant pas plus d'une erreur²⁷. Les résultats sont résumés dans le tableau 3.

Tableau 3. Pourcentage de réponses correctes aux questions Q1 et Q2 du OCM administré en fin de séance

	Groupes de contrôle	Groupes expérimentaux
Pourcentage de réponses correctes à Q1	39,58	38,57
Pourcentage de réponses correctes à Q2	64,58	58,57
Nombre de groupes	2	3
Nombre d'étudiants	48	70

Contrairement à Frank, nous ne trouvons pas de différence significative entre les groupes expérimentaux et les groupes classiques : la statistique du χ^2 pour l'hypothèse d'égalité du pourcentage de réponses correctes dans les deux types de groupes est égale à 0,012 ($p = 0,912$) pour la question Q1 et 0,433 ($p = 0,511$) pour la question Q2.

3.1.3. Conclusion sur les évaluations de fin de séance

L'ensemble des résultats obtenus sur les deux OCM conduit à la conclusion d'une absence de différence significative entre les étudiants des groupes expérimentaux et ceux des groupes classiques en termes de performance moyenne. Autrement dit, le fait d'avoir réalisé l'expérience ne semble pas avoir d'effet significatif sur le résultat obtenu lors des deux OCM administrés en fin de séance. Ce résultat est clairement en contradiction avec l'hypothèse H1 concernant l'impact pédagogique de court terme des expériences. On notera que, s'il est contraire aux observations de Frank [1997], il

²⁶ Par exemple, un test ANOVA donne une statistique F de Fisher égale à 0,029 ($p = 0,865$).

²⁷ Comme Frank [1997, p. 766], nous pensons que ce critère est raisonnable, notamment compte tenu de l'ambiguïté de certaines propositions.

est conforme à d'autres résultats obtenus dans la littérature (Cardell et al. [1996], Mullin et Sohan [1999], Yandell [1999]).

3.2. Résultats des évaluations réalisées trois mois plus tard (hypothèse H2)

L'analyse des résultats des évaluations réalisées trois mois après l'expérience nous permet d'évaluer la pertinence de l'hypothèse H2 concernant l'impact de « long terme » de l'expérimentation. Cette évaluation s'est opérée en deux temps : un test de connaissance générale comportant deux questions libres (QA : définition de la « tragédie des ressources communes » et QB : exemple de « tragédie des ressources communes ») puis le même OCM que celui administré juste après la séance, reprenant simplement les questions Q1 et Q2.

3.2.1. Le test de connaissance générale

Le tableau 4 donne les résultats moyens des deux catégories d'étudiants pour les questions QA et QB²⁸.

Tableau 4. Notes moyennes aux questions QA et QB lors de l'évaluation réalisée trois mois après la séance

	Groupes de contrôle	Groupes expérimentaux
Nombre de groupes	2	3
Nombre d'étudiants	47	64
Note moyenne (sur 4) pour la question QA (écart-type entre parenthèses)	0,808 (0,900)	1,843 (1,184)
Note moyenne (sur 4) pour la question QB (écart-type entre parenthèses)	1,085 (1,017)	1,515 (1,038)

Rappel : QA : Expliquez le plus précisément possible ce qu'on appelle « tragédie des ressources communes » ; QB : Donnez un exemple précis et détaillé de « tragédie des ressources communes ».

²⁸ En raison d'absences lors de cette seconde évaluation, nous avons moins d'étudiants que trois mois plus tôt : à savoir au total 111 (64 dans les groupes expérimentaux et 47 dans les groupes de contrôle) contre 118 (70 + 48) en janvier.

Les résultats sont saisissants. Les étudiants des groupes expérimentaux obtiennent des notes nettement plus élevées (plus du double en moyenne pour la question QA) que ceux des groupes classiques. Pour la question QA, la différence est significative au seuil de 1%, quel que soit le test statistique retenu ; par exemple, un test ANOVA donne une statistique F de Fisher égale à 24,746 ($p < 0,0001$). Concernant la question QB, la différence est également significative, quoique moins fortement : le test ANOVA donne une statistique de Fisher égale à 4,649, soit une différence significative au seuil de 5% ($p = 0,033$). Ainsi, les résultats obtenus à partir des questions QA et QB plaident clairement en faveur de l'hypothèse H2.

Ce résultat peut être confirmé et renforcé sur la base d'une analyse économétrique permettant de contrôler l'influence d'autres variables, en particulier concernant les caractéristiques individuelles des étudiants. Dans cette optique, nous avons collecté un certain nombre de données individuelles sur les étudiants impliqués dans l'étude. En raison de données manquantes sur 6 étudiants, nous n'avons que 105 observations. Bien entendu, la variable expliquée, *NOTE*, est la note obtenue par l'étudiant²⁹. Dans les variables explicatives, nous incluons l'appartenance à un groupe expérimental (variable indicatrice *EXP* prenant la valeur 1 si l'étudiant appartient à un groupe expérimental), le sexe (variable indicatrice *FEM* prenant la valeur 1 si l'étudiant est une fille), l'âge (variable *AGE*), la série du baccalauréat (variables indicatrices sur les trois séries : *ES*, *S* et *L*, prenant la valeur 1 lorsque l'étudiant a passé un bac *ES*, *S* ou *L*, respectivement) et la moyenne générale (sur 20) obtenue au baccalauréat (variable *MOYBAC*). Le tableau 5 donne les principales statistiques descriptives de l'échantillon.

Tableau 5. Statistiques descriptives

Variable	Global ($N = 105$)	Groupes expérimentaux ($N = 58$)	Groupes de contrôle ($N = 47$)
<i>NOTE</i> (QA)	1,40 (1,18)	1,88 (1,17)	0,81 (0,90)
<i>NOTE</i> (QB)	1,33 (1,06)	1,53 (1,06)	1,09 (1,02)
<i>EXP</i>	0,55	1	0
<i>FEM</i>	0,57	0,58	0,55

29. Nous nous intéressons ici aux deux questions QA et QB. Nous procédons donc à deux régressions, l'une expliquant la note à la question QA, l'autre expliquant la note à la question QB, le modèle ayant à chaque fois la même structure (les mêmes variables explicatives).

	Global ($N = 105$)	Groupes expérimentaux ($N = 58$)	Groupes de contrôle ($N = 47$)
<i>AGE</i>	19,4 (0,75)	19,39 (0,79)	19,40 (0,71)
<i>L</i>	0,17	0,16	0,19
<i>S</i>	0,50	0,52	0,47
<i>ES</i>	0,33	0,32	0,34
<i>MOYBAC</i>	14,44 (1,95)	14,56 (1,97)	14,27 (1,93)

Notes : (1) Les données manquantes concernent 6 étudiants appartenant à un groupe expérimental ; (2) en dehors des deux notes, les différences de moyenne entre les groupes expérimentaux et les groupes de contrôle ne sont significatives pour aucune des variables considérées.

Rappel : QA : Expliquez le plus précisément possible ce qu'on appelle « tragédie des ressources communes » ; QB : Donnez un exemple précis et détaillé de « tragédie des ressources communes ».

Notre modèle économétrique se présente donc de la manière suivante :

$$NOTE_i = \beta_0 + \beta_1 EXP_i + \beta_2 FEM_i + \beta_3 AGE_i + \beta_4 L_i + \beta_5 ES_i + \beta_6 MOYBAC_i + \varepsilon_i$$

où ε_i est le terme d'erreur pour l'étudiant i , β_0 la constante et β_1 à β_6 les coefficients à estimer. Notre hypothèse H2 prédit une influence positive de l'appartenance à un groupe expérimental sur le résultat de l'étudiant, soit $\beta_1 > 0$. Par ailleurs, si aucun effet spécifique n'est attendu sur les variables *FEM* ($\beta_2 = 0$) et *AGE* ($\beta_3 = 0$)³⁰, la série du bac et la moyenne générale obtenue au bac peuvent avoir une influence. En particulier, on peut s'attendre à des résultats plutôt moins bons chez les bacheliers littéraires ($\beta_4 < 0$)³¹, et, indépendamment de la série du bac, des résultats meilleurs de la part des meilleurs bacheliers ($\beta_6 > 0$).

Les résultats de l'estimation du modèle par la méthode des moindres carrés ordinaires (MCO) sont présentés dans le tableau 6.

30. Notons que, dans notre échantillon d'étudiants, un âge plus avancé ne correspond pas forcément à un redoublement. En effet, pour l'année 2004/2005 qui nous intéresse, les étudiants de 1^{re} année de l'Institut d'Études Politiques de Strasbourg étaient recrutés sur concours, soit directement après le bac (Bac + 0), soit un an après l'obtention du bac (Bac + 1). Par ailleurs, le redoublement de la 1^{re} année n'était pas de droit, mais ne pouvait être autorisé que sur décision spécifique du jury. Ainsi, l'échantillon ne comprend que cinq redoublants (répartis, comme les autres étudiants, dans les différents groupes suivant l'ordre alphabétique).

31. Notons que nous prenons la série scientifique comme série de référence et, par conséquent, c'est la variable indicatrice *S* qui est omise dans les régressions.

Tableau 6. Résultats des régressions

Variables indépendantes	Variable expliquée : NOTE (QA)	Variable expliquée : NOTE (QB)
Constante	0,571 (0,161)	3,895 (1,102)
EXP	1,037*** (4,949)	0,445** (2,129)
FEM	-0,032 (-0,144)	0,109 (0,497)
AGE	-0,051 (-0,319)	-0,150 (-0,949)
L	-0,235 (-0,783)	0,079 (0,266)
ES	0,041 (0,177)	-0,067 (-0,287)
MOYBAC	0,089 (1,429)	0,004 (0,065)
N	105	105
R ² ajusté	0,193	0,004

Note : t-statistique entre parenthèses : (*) significatif à 10 %, (**) significatif à 5 %, (***) significatif à 1 %.

Rappel : QA : Expliquez le plus précisément possible ce qu'on appelle « tragédie des ressources communes » ; QB : Donnez un exemple précis et détaillé de « tragédie des ressources communes ».

Sans surprise, l'effet de l'appartenance à un groupe expérimental apparaît hautement significatif. Ainsi, la prise en compte des autres variables ne change en rien la conclusion précédente : les étudiants des groupes expérimentaux obtiennent des résultats nettement meilleurs que ceux des groupes classiques dans l'évaluation réalisée trois mois après la séance et portant sur les deux questions QA et QB. Plus précisément, l'estimation de β_1 , le coefficient pour la variable EXP, est de 1,037 pour la question QA ($t = 4,949$; $p < 0,0001$) et 0,445 pour la question QB ($t = 2,129$; $p = 0,036$) ; ainsi, toutes choses égales par ailleurs, appartenir à un groupe expérimental améliore la note d'environ 1,04 point (sur 5) à la question QA et de 0,45 point (sur 5) à la question QB32.

On peut ajouter deux remarques. En premier lieu, les autres variables n'ont pas d'effet significatif. Si l'absence d'effet des variables FEM et AGE était attendue, on peut être un peu plus surpris pour ce qui concerne les variables L et MOYBAC. Notons toutefois que, dans la régression de la note à la question QA, la variable MOYBAC a, comme attendu, un effet légèrement positif, même si cet effet n'est pas significatif au seuil de 10 % ($t = 1,429$; $p = 0,156$). En second lieu, on notera la faible capacité prédictive

32. Sans surprise, compte tenu de la non-significativité des autres variables explicatives du modèle, la prime d'appartenance à un groupe expérimental correspond pratiquement à l'écart des moyennes entre les deux catégories de groupes, écart égal à 1,07 (1,88 - 0,81) pour la question QA et à 0,14 (1,53 - 1,09) pour la question QB.

de la seconde régression, celle expliquant la note à la question QB (R^2 ajusté = 0,004).

3.2.2. Le QCM

Les résultats obtenus dans le cadre du QCM vont dans le même sens que ceux obtenus à partir du test de connaissance générale, mais les différences ne sont pas significatives. Le tableau 7 récapitule les résultats obtenus sur les questions Q1 et Q2 lors du QCM administré trois mois après la séance.

Tableau 7. Pourcentage de réponses correctes aux questions Q1 et Q2 du QCM administré trois mois après la séance

	Groupes de contrôle	Groupes expérimentaux
Pourcentage de réponses correctes à Q1	38,30	43,75
Pourcentage de réponses correctes à Q2	40,43	45,31
Nombre de groupes	2	3
Nombre d'étudiants	47	64

Les étudiants des groupes expérimentaux réussissent un tout petit peu mieux que leurs homologues des groupes classiques dans le QCM administré trois mois après la séance (ce qui n'était pas le cas juste après la séance). Cependant, la différence entre les deux catégories n'est pas significative, ni pour Q1 ($\chi^2 = 0,332$; $p = 0,565$), ni pour Q2 ($\chi^2 = 0,264$; $p = 0,608$)33.

Les différences observées dans le cadre du QCM sont faibles et non significatives alors qu'elles sont fortes et très significatives sur les questions ouvertes. Selon nous, cela n'est guère surprenant. Cela suggère simplement que les questions ouvertes permettent une mesure beaucoup plus fine de l'apprentissage des étudiants que le format QCM. En effet, face aux questions QA et QB, les étudiants devaient restituer une compréhension globale des concepts alors que, face aux questions Q1 et Q2, ils étaient guidés par les différentes propositions et, par conséquent, « aidés » par le format des questions. Dans le cadre de l'évaluation réalisée trois mois après la séance, il est clair que les réponses aux deux questions ouvertes sont plus représentatives de la capacité des étudiants à se remémorer les concepts que les deux questions fermées du QCM.

33. On notera que, pour la question Q1, les étudiants des groupes expérimentaux ont légèrement amélioré leur performance par rapport au premier QCM (43,75 % de réponses correctes dans le second QCM contre 38,37 % dans le premier), au contraire des étudiants des groupes de contrôle dont la performance d'ensemble a, au contraire, très légèrement baissé (38,30 % de réponses correctes dans le second QCM contre 39,58 % dans le premier).

3.2.3. Conclusion sur les évaluations réalisées trois mois après la séance

Au total, nos résultats plaident pour une différence entre les deux catégories d'étudiants en faveur des groupes « expérimentaux ». En particulier, sur la question libre de définition du concept de « tragédie des ressources communes » (question Q4), ils obtiennent une note moyenne plus de deux fois supérieure (1,84 contre 0,81), une différence hautement significative, que l'on tienne compte ou non des autres variables pouvant intervenir dans la performance des étudiants.

Conclusion

Au final, les résultats de notre étude tendent à invalider l'hypothèse H1, mais à confirmer l'hypothèse H2. Ce serait donc bien dans l'apprentissage de long terme que résiderait principalement l'efficacité pédagogique de l'expérimentation dans les cours d'économie. Cela rejoint certains principes importants des sciences de l'éducation, notamment la « théorie de l'apprentissage expérimental » (Kolb [1984]). Selon cette théorie, les étudiants apprennent mieux lorsque l'enseignement prend une forme expérimentale. Avec les jeux pédagogiques, l'étudiant est « acteur » et non plus simple « spectateur » du cours. Il en résulte des modifications à la fois en termes de motivation et en termes de mode d'apprentissage, les deux principales raisons généralement invoquées pour justifier l'apprentissage plus « profond », plus « durable », produit par la démarche expérimentale. En premier lieu, les situations dans lesquelles les élèves sont plutôt « acteurs » alimentent une implication personnelle de l'élève et, par suite, une motivation intrinsèque, la motivation extrinsèque correspondant à l'élève ordinaire, qui est un peu « spectateur » du cours (Lieuury et de la Haye [2004, p. 114]). En second lieu, l'expérimentation d'un concept implique un apprentissage plus profond. Alors qu'un enseignement standard, incluant textes et images dans le cadre d'un cours magistral, induit un double codage des informations (verbal + imagé), l'expérimentation permet en plus un codage moteur, ce qui conduit au final à un triple codage (verbal + imagé + moteur) des informations (Lieuury et de la Haye [2004, p. 32]). Dans une perspective de mémoire de long terme, les informations sont d'autant plus ancrées que le codage est « épais ». Ainsi, apprendre par l'action est plus efficace qu'apprendre par la parole : les élèves oublient moins facilement et moins vite ce qu'ils ont eux-mêmes expérimenté.

Au final, nos résultats plaident naturellement plutôt pour l'inclusion de jeux pédagogiques dans les cursus en arguant du fait que, même si les résultats des étudiants n'apparaissent pas sensiblement affectés à court terme par cette nouvelle méthode pédagogique, la rétention du savoir à long terme pourrait en être largement améliorée. Cependant, nous n'avons pas cherché à évaluer, dans le cadre de cette recherche, les coûts liés à l'utilisation de la méthode expérimentale. Or ceux-ci peuvent s'avérer signi-

ficatifs : temps pris sur la présentation « traditionnelle » des concepts, temps de préparation supplémentaire pour l'enseignant, difficultés d'organisation, etc. (cf. Eber [2003]). Ainsi, un prolongement de ce travail serait d'adapter le dispositif pour évaluer de manière globale les bénéfices et les inconvénients de la méthode expérimentale par rapport à la méthode classique.

Références bibliographiques

- BELL R. et DELEMEESTER G. [1999], The Double Oral Auction : Is It an Effective Teaching Tool ?, in *Proceedings, Economics and the Classroom Conference*, Prentice-Hall, p. 12-32.
- BRAUER J. et DELEMEESTER G. [2001], Games Economists Play : A Survey of Non-Computerized Classroom-Games for College Economics, *Journal of Economic Surveys*, 15, p. 221-236.
- CARDELL S., FORT R., JOERDING W., INABA F., LAMOREAUX D., ROSENMAN R., STROMSDORFER E. et BARTLETT R. [1996], Laboratory-Based Experimental and Demonstration Initiatives in Teaching Undergraduate Economics, *American Economic Review*, 86, p. 454-459.
- DICKIE M. [2006], Do Classroom Experiments Increase Learning in Introductory Microeconomics ?, *Journal of Economic Education*, 37, p. 267-288.
- EBER N. [2003], Jeux pédagogiques : vers un nouvel enseignement de la science économique, *Revue d'Economie Politique*, 113, p. 485-521.
- EBER N. [2006], L'économie expérimentale comme outil pédagogique, Site SES-ENS de l'ENS - Lettres et Sciences Humaines de Lyon (<http://ses.ens-lsh.fr>).
- EBER N. et WILLINGER M. [2005], *L'économie expérimentale*, La Découverte, collection Repères (n° 423).
- EMERSON T. et TAYLOR B. [2004], Comparing Student Achievement across Experimental and Lecture-Oriented Sections of a Principles of Microeconomics Course, *Southern Economic Journal*, 70, p. 672-693.
- FRANK B. [1997], The Impact of Classroom Experiments on the Learning of Economics : An Empirical Investigation, *Economic Inquiry*, 35, p. 763-769.
- GREMMEN H. et POTTERS J. [1997], Assessing the Efficacy of Gaming in Economic Education, *Journal of Economic Education*, 28, p. 291-303.
- HOLT C. [1999], Teaching Economics with Classroom Experiments : A Symposium, *Southern Economic Journal*, 65, p. 603-610.
- HOLT C. et LAURY S. [1997], Classroom Games : Voluntary Provision of a Public Good, *Journal of Economic Perspectives*, 11, p. 209-215.
- KOLBD. [1984], *Experiential Learning*, Prentice-Hall.
- LIEURY A. et DE LA HAYE F. [2004], *Psychologie cognitive de l'éducation*, Dunod.
- MULLIND. et SOHAN G. [1999], Benefit Assessment of Classroom Experimental Economics, Department of Economics and Geography, USAF Academy.
- VANDELL D. [1999], Effects of Integration and Classroom Experiments on Student Learning and Satisfaction, in *Proceedings, Economics and the Classroom Conference*, Idaho State University.