



HAL
open science

Une méthodologie d'évaluation d'un dispositif d'aide aux primoentrants à l'université

Isabelle Bournaud, Patrick Pamphile

► To cite this version:

Isabelle Bournaud, Patrick Pamphile. Une méthodologie d'évaluation d'un dispositif d'aide aux primoentrants à l'université. 2020. hal-03089701

HAL Id: hal-03089701

<https://hal.science/hal-03089701>

Preprint submitted on 28 Dec 2020

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.

Copyright

Une méthodologie d'évaluation d'un dispositif d'aide aux primo-entrants à l'université

Isabelle Bournaud, Groupe Didasco - UR EST - UFR Sciences, Université Paris Saclay, isabelle.bournaud@universite-paris-saclay.fr

Patrick Pamphile, Laboratoire de Mathématiques d'Orsay (LMO), Université Paris Saclay, patrick.pamphile@universite-paris-saclay.fr

Résumé

Evaluer un dispositif pédagogique est une préoccupation importante des universités françaises aujourd'hui. D'une part parce que ces dispositifs s'inscrivent dans la lutte contre l'échec massif en premier cycle universitaire, ils sont considérés comme un investissement ; et d'autre part dans le cadre de l'amélioration de la qualité des formations.

Cet article présente une méthodologie d'évaluation d'un dispositif basée sur trois critères d'évaluation identifiés dans la littérature : l'efficacité pédagogique, l'efficacité et l'équité pédagogique. Ces critères sont évalués à partir d'indicateurs calculés sur la perception de maîtrise des thèmes travaillés dans le dispositif, rapportée par les étudiants. Cette méthodologie a été utilisée pour évaluer deux dispositifs mis en place à destination de primo-entrants en 1^{ère} année universitaire dans deux composantes de l'Université Paris-Saclay.

Abstract

Evaluating a support program has become a major concern for French universities. On the one hand because these programs are part of the struggle against massive failure in undergraduate studies, they are considered as an investment; and on the other hand to improve the quality of training.

This article presents an evaluation methodology based on three evaluation criteria identified in the literature : pedagogical effectiveness, efficiency and pedagogical equity. These criteria are evaluated on the basis of indicators calculated on the students' perception of their level of proficiency in the skills addressed in the curriculum, before and after the program is completed. This methodology has been used to evaluate two support program for newcomers in their first year at the Université Paris-Saclay.

Mots clés

Evaluation de dispositif pédagogique, efficacité pédagogique, efficience, équité pédagogique, perception de maîtrise, 1^{ère} année universitaire

Keywords

Support program assessment, pedagogical effectiveness, efficiency and pedagogical equity, self-evaluation of skill proficiency, first year undergraduate

1. Introduction

Avec la démocratisation de l'accès à l'enseignement supérieur, la mise en place de dispositifs pédagogiques visant à améliorer la réussite des étudiants est devenue une préoccupation importante des universités. Compte tenu des taux d'échec ou d'abandon importants en 1^{ère} année d'université, ces dispositifs sont considérés d'une part comme un investissement et d'autre part comme faisant partie de la formation en premier cycle. L'évaluation du dispositif est alors indispensable pour le valoriser, pour rendre compte aux différentes parties prenantes (étudiants, enseignants, institution l'ayant financé, ...) et également afin de l'améliorer (Bernard, 2011; Salmon et al., 2009). En outre, lorsque les actions sont le fruit de décisions gouvernementales, telles que le Plan Réussite en Licence¹ (PRL) en France entre 2007 et 2012 ou loi dite ORE², et le Plan Etudiants en 2018, l'élaboration d'un bilan est indispensable.

Différents modèles ont été proposés pour l'évaluation d'une formation. Parmi ceux-ci, le modèle de Kirpatrick est largement utilisée (Monnot, 2014). Dans le contexte universitaire, les travaux menés mettent en évidence que l'évaluation d'un dispositif pédagogique est complexe et difficile à analyser (Michaut, 2003; Perret, 2015; Salmon et al., 2009). Les difficultés résident notamment dans le choix des critères d'évaluation du dispositif (validation du semestre ou de l'année, développement de méthodes de travail, affiliation de l'étudiant), dans le recueil de données représentatives, dans le choix de l'outil de recueil (questionnaire à questions fermées ou ouvertes, entretiens, ...), dans le choix des indicateurs à analyser

¹ PRL : Plan Réussite en Licence <http://www.enseignementsup-recherche.gouv.fr/cid20651/plan-pour-la-reussite-en-licence-730-millions-d-euros-d-ici-2012.html>

² LOI n° 2018-166 du 8 mars 2018 relative à l'Orientation et à la Réussite des Etudiants (ORE)

(évaluation quantitative, qualitative,...), ou encore dans le moment où l'évaluation est réalisée (à chaud ou à froid). De surcroit, une difficulté inhérente aux recherches en éducation est que la validité de l'évaluation peut être remise en question : les très nombreuses variables qui interviennent dans la réussite étudiante peuvent varier d'une année à l'autre, et ne sont pas toutes considérées simultanément (Sanchez et Monod-Ansaldi, 2015).

Nous aborderons ainsi trois questions recherche relatives à l'évaluation d'un dispositif pédagogique, :

- quels critères ? quelles sont les finalités du dispositif ?
- quelles données recueillir ? et comment ?
- quels indicateurs construire sur les données pour évaluer les critères à partir des données ?

Le cadre théorique de ce travail est présenté dans la partie 2. La méthodologie d'évaluation proposée est exposée dans la partie 3. Pour finir, les atouts et limites de cette méthodologie sont discutés dans la partie 4.

2. Cadre théorique

Evaluer un dispositif de formation permet de l'améliorer, de le valoriser et de rendre compte aux différentes parties prenantes (étudiants, enseignants, institution l'ayant financé, ...) (Depover, et al., 2018 ; Bernard, 2011 ; Salmon et al., 2009).

Dans le contexte universitaire français, avant la mise en place institutionnelle du PRL entre 2007 et 2012, puis de la loi ORE et du Plan Etudiants en 2018, les travaux sur l'évaluation des dispositifs pédagogiques d'aide aux étudiants n'avaient que peu donné lieu à une évaluation (Michaut, 2003). En effet, les équipes impliquées dans les actions pédagogiques d'accompagnement concentraient leurs actions sur l'aide aux étudiants plus que sur l'évaluation du dispositif (Morlaix et Perret, 2013 ; Salmon et al., 2019). En outre, mettre en place une évaluation requiert l'obtention de données qui ne sont pas toujours aisément disponibles (par exemple les résultats académiques de l'année) et une compétence dans l'analyse de ces données. Perret et Morlaix en 2014 (p.188) indiquent que « la recherche sur la mesure des dispositifs pédagogiques à l'université reste balbutiante ». L'institutionnalisation de dispositifs d'aide aux primo arrivants a alors rendu leur évaluation plus systématique.

Des modèles d'évaluation ont d'abord été proposés dans le cadre de la formation professionnelle, et plus récemment dans le contexte scolaire puis universitaire.

2.1.1. Quels modèles ? Quels critères ?

Dans les années 60, Kirkpatrick a mis au point un modèle d'évaluation hiérarchique d'une formation, qui repose sur 4 niveaux : 1-réaction (les participants ont-ils apprécié les modalités pédagogiques, la formation ?) ; 2-apprentissage (les participants ont-ils fait des apprentissages ?) ; 3-comportement (les participants ont-ils transféré les apprentissages réalisés en situation professionnelle ?) ; 4-résultats (la formation a-t-elle eu un effet sur l'organisation ?). Plus le niveau est élevé, plus il fournit de l'information qui revêt de l'importance pour l'entreprise. Il ne s'agit donc pas de choisir d'évaluer selon un de ces quatre critères, mais plutôt d'évaluer au regard des quatre critères. Ce modèle, ou une déclinaison de celui-ci (Dunberry et Péchard, 2007 ; Praslova, 2010), bien que critiqué notamment sur la relation causale entre les niveaux, est largement utilisé dans les entreprises françaises pour l'évaluation de formations professionnelles (Monnot, 2014). Ce modèle a également désormais pleinement intégré le champ de la pédagogie universitaire. Ces différents modèles ont le mérite d'amener à expliciter précisément les finalités de l'action, ce qui constitue un pas important dans l'évaluation (Dunberry et Péchard, 2007). Le lecteur intéressé par ces différents modèles pourra se référer à Tamkin et al. (2002).

Selon Gérard (2001), certains critères utilisés pour évaluer une formation professionnelle peuvent également servir à évaluer les systèmes éducatifs. Ainsi, les experts s'accordent pour articuler la problématique de l'évaluation des systèmes éducatifs autour des concepts-clés de l'efficacité et l'efficience, auxquels s'ajoute celui d'équité, dont Sall et De Ketele (1997) montrent l'importance.

- **L'efficacité** d'une action de formation peut être évaluée soit par une évaluation interne, soit par une évaluation externe. L'évaluation interne vise à évaluer les acquis d'apprentissages. Plus précisément, il s'agit de vérifier que les objectifs de la formation ont été atteints, que les participants ont acquis, à la fin de la formation, les compétences visées. On parle ainsi de *l'efficacité pédagogique* de la formation (Gérard, 2001 ; Sall et De Ketele, 1997). L'évaluation externe quant-à-elle a deux visées : d'une part évaluer le transfert des apprentissages (Les acquis de la formation sont-ils appliqués sur le terrain, dans d'autres cadres de la formation universitaire ? Les participants, une fois revenus sur leur poste de travail, mettent-ils en œuvre les compétences acquises lors de la

formation ?), et d'autre part évaluer l'impact de la formation sur l'entreprise (Les compétences acquises lors de la formation permettent-elles de faire évoluer l'organisation ?). L'évaluation externe vise ainsi à évaluer l'impact de la formation sur le terrain. Notons que sur le terrain, il est difficile d'isoler ce qui est dû uniquement à la formation. Pourtant, comme le souligne Gérard (2001), c'est ce point qui intéresse véritablement l'entreprise dans le cadre de la formation professionnelle.

- **L'équité** d'une formation, selon Sall et De Ketele (1997), se décline en cinq volets : équité en matière d'accès, de confort, de production, d'accomplissement et équité pédagogique (réduction de l'écart entre les plus performants et les moins performants entre le début et la fin de l'action pédagogique). Il n'est pas inutile d'indiquer que l'équité et l'efficacité ne sont pas nécessairement liés (Gérard, 2001). En effet, les apprenants peuvent avoir atteint les objectifs de la formation sans pour autant que l'écart entre les apprenants ait été réduit, notamment s'il y a une forte hétérogénéité de maîtrise au départ.
- Enfin, **l'efficience** est le rapport entre les résultats effectivement obtenus à la fin de l'action et la nature et le volume des moyens mis en œuvre pour réaliser l'action ; il s'agit donc de l'efficacité rapportée aux moyens utilisés. Elle se décline également en efficience interne, et efficience externe (Sall et De Ketele, 1997).

Il est intéressant de noter que dans leurs travaux sur la mesure de la qualité, Bouchard et Plante (2002) proposent un modèle basé sur neuf qualités transversales, parmi lesquelles on retrouve la pertinence, l'efficacité et l'efficience.

Pour Lessard et Meirieu (2005), cité dans De Ketele et Gerard (2007) « La norme ultime du pilotage réussi restera toujours l'efficacité en termes de performances des élèves ainsi que l'équité pédagogique. Il ne s'agit pas seulement d'être efficace, mais aussi équitable. Il s'agit non seulement que l'ensemble du groupe d'élèves progresse, mais aussi que le fossé qui, au départ, sépare les élèves les plus faibles des plus forts ne se creuse pas et, si possible, se réduise. »

2.1.2. Quelles données et quels indicateurs à considérer ?

Le groupe de recherche AdAPTE (Actions d'Accompagnement Pédagogique, leur Typologie et leur Evaluation) a proposé une grille d'analyse des actions d'accompagnement pédagogiques des primo-arrivants dans l'enseignement supérieur (Salmon et al., 2009). Six éléments constitutifs de l'évaluation sont identifiés : le commanditaire, les buts poursuivis, les

données collectées, les outils utilisés pour la collecte, les acteurs interrogés, le moment de l'évaluation.

La méthode proposée par Gérard (2001) consiste à évaluer l'efficacité pédagogique d'une formation via une auto-évaluation par les apprenants de leur niveau de maîtrise pour chacune des compétences générales et spécifiques avant et après la formation. Mesurer l'écart avant-après ne correspond pas une évaluation des acquis d'apprentissage par l'enseignant, mais constitue néanmoins un indicateur de mesure de l'efficacité pédagogique. Gérard (2001) suggère de considérer en plus des gains bruts (différence de niveau entre après et avant), les « gains relatifs » qui correspondent à ce qui aurait pu être gagné grâce à la formation. Concernant le moment où se déroule l'évaluation, l'auteur considère que les auto-évaluations du niveau de maîtrise avant et après la formation peuvent être réalisées toutes les deux à la fin de la formation.

Michaut (2003) propose de mesurer l'efficacité d'un dispositif pédagogique au regard de l'amélioration de la réussite à l'année, de la diminution des abandons, ou encore si le dispositif favorise l'intégration universitaire ou le développement des compétences des étudiants. Il met cependant l'accent sur la difficulté de mener une telle évaluation :

- l'écart entre les taux de réussite (par exemple) des étudiants ayant suivi le dispositif et ceux ne l'ayant pas suivi peut être biaisé par un « effet groupe ». Pour l'éviter, il est nécessaire d'avoir une forte concordance entre les deux groupes, ce qui est rarement le cas notamment lorsque le dispositif est optionnel, ou obligatoire pour une sous-population spécifique ;
- la présence à l'université n'étant pas toujours obligatoire, la fréquentation du dispositif peut ne pas être un indicateur pertinent.

Dans le cadre du PRL, différents dispositifs à destination des primo-entrants ont été mis en œuvre, variant selon les universités et au sein même des universités³. Morlaix et Perret (2013) présentent une analyse des effets du PRL, à l'université de Bourgogne, qui repose sur une analyse de la validation du diplôme à l'aide de modèles de régression logistique. Les facteurs considérés sont les caractéristiques sociodémographiques des étudiants, la filière, l'année

³ Rapport du Comité de suivi de la licence et de la licence professionnelle. Année 2010-2011, CPU/Paris, 2011, p.49 http://www.cpu.fr/wp-content/uploads/2013/07/rapport_CSL_2010_2011.pdf

universitaire, les actions PRL suivies ; l'indicateur d'évaluation de l'efficacité du dispositif est la réussite au semestre 1 (Admis, Ajourné, Abandon).

Ces différents travaux se focalisent principalement sur la dimension de l'efficacité du dispositif pédagogique, qui est évaluée au regard des résultats académiques au semestre ou à l'année (Michaut, 2003 ; Perret, 2015 ; Salmon et al., 2009).

Une autre approche consiste à se baser sur les avis des étudiants sur l'impact du dispositif. C'est ainsi que le Service Universitaire de pédagogie de l'Université Joseph Fourier de Grenoble a évalué, entre 2010 et 2014, deux approches mises en place pour favoriser le développement des méthodes de travail des étudiants, dans le cadre du PRL (Cosnefroy et al., 2014). Les auteurs se sont basés sur les avis des étudiants sur l'évolution de leurs pratiques et sur le dispositif, recueillis via des questionnaires d'évaluation administrés à la fin du dispositif. Ces questionnaires comportaient une vingtaine de questions fermées (réponses portées sur une échelle de Lickert à 4 modalités), une question d'appréciation globale avec une échelle de 6 modalités, et 3 questions ouvertes (point forts, points faibles, points à améliorer). Les auteurs utilisent comme indicateur pour évaluer l'impact du dispositif le fait que les étudiants déclarent avoir fait évoluer leurs pratiques.

Perrault, Brassart et Dubus (2010) utilisent pour leur part le sentiment d'efficacité personnelle comme indicateur de la qualité d'une formation pour des futurs enseignants, couplé à une auto-évaluation des participants de leurs activités et également une évaluation des apprentissages par les formateurs.

Figari et Remaud (2014) ont proposé le modèle ICP (Induit – Construit – Produit) dans lequel on retrouve les deux critères d'évaluation : efficacité et efficience, pour les volets pédagogique et économique, auxquels s'ajoute la pertinence qui concerne le volet politique de la mise en place d'un dispositif pédagogique à l'université. La pertinence mesure la conformité entre le dispositif et les objectifs visés par l'établissement. Par exemple, un dispositif à destination des étudiants ayant obtenu une mention TB (Très Bien) au baccalauréat est-il pertinent pour favoriser la réussite de tous les primo-entrants en 1^{ère} année universitaire ?

Pour résumer, dans les travaux existants sur l'évaluation d'un dispositif pédagogique, différentes données sont considérées : des données sur le profil des étudiants et leur taux de participation ; des données concernant leur degré de satisfaction par rapport à l'atteinte des objectifs annoncés, la qualité du dispositif ainsi que la réponse à leurs attentes ; des données

sur le ressenti des étudiants (utile/pertinent, rassurant, motivant, soutenant, incitant, ...). Par ailleurs, des données sur les effets du dispositif sont également prises en compte : la réussite aux examens et l'évolution du comportement des étudiants. Toutefois, dans la majorité des cas, l'indicateur utilisé pour évaluer l'efficacité du dispositif est la réussite aux examens. Or, si les dispositifs pédagogiques visent à améliorer la réussite, cette dernière dépend de nombreux autres facteurs : contexte d'études, connaissances initiales, pratiques d'étude, modalités de contrôles des connaissances (Berthaud, 2017 ; Michaut et Romainville, 2012 ; Paivandi, 2015). Par ailleurs, les dispositifs peuvent viser des objectifs intermédiaires à la réussite : le développement de stratégies d'étude efficaces, l'épanouissement des étudiants, leur persévérance, une meilleure affiliation ... (Gilibert et Gillet, 2010). Ainsi, des indicateurs permettant de mesurer des changements chez les apprenants doivent également être proposés. Enfin, les indicateurs doivent fournir des pistes d'amélioration du dispositif en termes d'équité pédagogique et d'efficacité.

3. Méthodologie d'évaluation proposée

Au vu de notre revue de littérature, nous avons retenu trois critères pour évaluer un dispositif pédagogique : l'efficacité pédagogique, l'efficacité et l'équité pédagogique.

Pour évaluer ces critères, nous utilisons différents indicateurs qui sont calculés, à l'instar de Gérard (2003) et Cosnefroy et al., (2014), à partir de la perception de maîtrise des étudiants des thèmes travaillés dans le dispositif, et de l'évolution de celle-ci avant et après le dispositif. La perception de maîtrise est auto-rapportée par les étudiants, à la fin du dispositif.

Cette méthodologie a été utilisée pour évaluer deux dispositifs à destination de primo-entrants dans des composantes de l'Université Paris-Saclay. Le lecteur intéressé trouvera la description d'un des dispositifs dans (Bournaud et Pamphile, 2020)

3.1. Données collectées

Il a été demandé aux étudiants d'exprimer leur perception de maîtrise sur différents items, avant et après le dispositif. Les items ont été élaborés au regard des thèmes travaillés dans les dispositifs considérés. Par exemple, pour le thème « Gérer son temps », un item proposé est « Je planifie mon travail sur la semaine », pour le thème « Prendre des notes » un item proposé est « Je suis capable d'extraire les informations essentielles d'un document ou d'un cours ».

Une échelle mixte de perception de maîtrise a été utilisée (voir tableau 1 ci-dessous). L'échelle numérique permet de faire des calculs ; on utilise des valeurs négatives et des valeurs positives pour traduire une perception négative ou positive. L'échelle qualitative permet à l'étudiant de se positionner.

Tableau 1: Exemple d'item

Q3-Je planifie mon travail sur la semaine										
	-5 Non, pas du tout	-4	-3	-2	-1 Non, pas vraiment	1 Oui, plutôt	2	3	4	5 Oui, tout à fait
Avant										
Après										

Enfin, il s'agit d'études longitudinales : les données sont collectées tous les ans (2 années à ce jour) afin de pouvoir mesurer l'évolution de l'évaluation des différents critères considérés.

3.2. Indicateurs utilisés

A partir de la perception de maîtrise avant et après le dispositif, rapportée par les étudiants, nous calculons différents indicateurs.

3.2.1. Equité pédagogique

Pour chacun des items, on calcule la moyenne et le coefficient de variation des perceptions de maîtrise, avant et après le dispositif (voir Tableau 2 ci-après).

Tableau 2: Indicateur de tendance et d'hétérogénéité.

Moyenne	\bar{x}	C'est un indicateur de <i>tendance</i> des réponses.
Coefficient de variation	$\frac{\sigma}{\bar{x}}$	σ est l'écart-type des réponses. Noté en pourcentage, le coefficient de variation est un indicateur <i>d'hétérogénéité</i> des réponses.

La moyenne est un indicateur de tendance centrale, le coefficient de variation est un indicateur de dispersion. Nous avons retenu comme indicateur le coefficient de variation car il permet de mesurer l'hétérogénéité des réponses, et par conséquent **l'équité pédagogique** du dispositif.

3.2.2. Efficacité et efficience

Par ailleurs, pour chacun des items on calcule deux indicateurs : le gain brut moyen et le gain relatif (voir Tableau 3 ci-après) :

Tableau 3: Gain brut et gain relatif

Gain brut moyen	$\bar{x}_{\text{après}} - \bar{x}_{\text{avant}}$	Correspond à ce qui a été « gagné » au cours de l'action
Gain relatif	$\Delta = \frac{\bar{x}_{\text{après}} - \bar{x}_{\text{avant}}}{x_{\text{max}} - \bar{x}_{\text{avant}}}$	C'est le rapport entre les moyennes de ce qui « a été gagné » et de ce qui « aurait pu être gagné » au cours de l'action (x_{max} étant la valeur maximale à atteindre).

Pour illustrer les notions de gain brut et de gain relatif, considérons l'exemple de réponses de quatre groupes d'étudiants (voir Tableau 4 ci-dessous).

Tableau 4: Exemple de gain brut et relatif

Groupe	\bar{x}_{avant}	$\bar{x}_{\text{après}}$	Gain brut	Gain relatif
A	-5	-4	1	10%
B	4	5	1	100%
C	-1	1	2	33%
D	-4	1	5	56%

Considérons les groupes A et B de cet exemple : on voit que le gain brut ne permet pas de discriminer ces deux groupes (le gain brut est de 1 pour les deux groupes) alors que les niveaux atteints ne sont pas les mêmes : le groupe A est arrivé à -4 alors que le groupe B est arrivé à 5. On utilise alors le gain relatif. Le gain relatif prend en compte ce qui aurait pu être gagné (c'est-à-dire l'écart entre le niveau de départ et le niveau maximum). Les groupes A et B n'ont effectivement pas le même gain relatif (respectivement 10% et 100%). Considérons les groupes C et D : ils ont atteint le même niveau après l'action ($x_{\text{après}} = 1$), mais n'étaient pas au même niveau avant (-1 pour le groupe C, -4 pour le groupe D). Les gains relatifs sont alors de 33% pour le groupe C et 56% pour le groupe D.

Ainsi, si le gain brut nous sert à évaluer **l'efficacité pédagogique** d'une action pédagogique, le gain relatif nous sert à évaluer son **efficience**. A l'instar de Gérard (Gérard, 2003), nous utiliserons le seuil de significativité de 30% : l'action est considérée comme efficiente si le gain relatif est supérieur à 30%.

Il n'est pas inutile de noter que le gain relatif est un rapport entre la moyenne du gain brut et la moyenne du gain possible, et non une moyenne des rapports des deux gains pour chaque étudiant. En effet, si des étudiants estimaient être au niveau maximal avant et après l'action, alors ils ont un gain brut individuel nul ($x_{\text{max}} - x_{\text{avant}} = 0$) et il n'est pas possible de calculer un gain relatif pour ces étudiants.

Notons que les données étant calculées sur un échantillon d'étudiants, il convient de valider les résultats à l'aide de tests statistiques (Saporta, 2006).

3.3. Illustration de l'analyse des indicateurs

Afin d'illustrer l'utilisation des indicateurs considérons le tableau 5 qui fournit deux items de l'évaluation d'un des dispositifs.

Tableau 5: Exemple d'indicateurs pour deux items

		Moyenne	Coefficient de variation	Gain relatif
Q1-Je connais les personnes ressources et leur rôle dans ma formation	Avant	-1	245,0%	46,0%
	Après	1,8	95,9%	
Q10-En général avant un contrôle, je sais quoi faire pour avoir une bonne note	Avant	1,3	131,7%	23,9%
	Après	2,3	72,9%	

L'analyse des indicateurs pour l'item Q1 du tableau 5 fournit les éléments suivants :

- **Gain brut moyen** : on constate une amélioration en moyenne du niveau de maîtrise perçu : on passe en moyenne de -1 avant à 1,8 après (gain brut moyen de 2,8). On peut considérer que l'action a été efficace. On valide cette décision à l'aide d'un test des rangs signés de Wilcoxon (Saporta, 2006). A titre d'exemple, ici l'écart avant-après est statistiquement significatif ($p\text{-value} < 0,001$).
- **Coefficient de variation** du niveau de maîtrise : il a nettement diminué entre avant (245,0%) et après (95,9%). Cela signifie que les réponses après sont moins hétérogènes qu'avant. Ce résultat met en avant l'équité pédagogique de cette action.
- **Gain relatif** : celui-ci étant supérieur à 30%, on considère que l'action a été efficace. Pour tester l'écart à 30% on utilise un test Z (Saporta, 2006). A titre d'exemple, ici l'écart avant-après est statistiquement significatif ($p\text{-value} < 0,001$).

Considérons maintenant l'item Q10 du Tableau 5 ci-dessus, et analysons les indicateurs :

- **Gain brut moyen** : la perception de maîtrise augmente puisque la moyenne passe de 1,3 avant les ateliers à 2,3 après : l'action a donc été efficace.
- **Coefficient de variation** : il a diminué entre avant (131,7%) et après (72,9%), ce qui met en avant l'équité pédagogique de cette action.
- **Gain relatif** : le gain relatif est de 23,9% (<30%), donc on ne peut pas dire que l'action a été efficace. Une explication de ce résultat que l'on pourrait avancer est que certains étudiants ont considéré qu'ils maîtrisaient cet élément avant l'action et ont pris conscience après les premières évaluations des apprentissages que ce n'était pas le cas. Quoiqu'il en soit, cette valeur du gain relatif inférieure à 30% peut être considérée comme un indicateur d'une piste d'amélioration pour l'action pédagogique.

4. Discussion : atouts et limites de la méthodologie proposée

Au regard de la revue de littérature, nous proposons dans ce travail de considérer les trois critères suivants pour évaluer un dispositif pédagogique : l'efficacité pédagogique, l'équité pédagogique et l'efficience. Pour évaluer ces critères, nous nous basons sur la perception du niveau de maîtrise rapportée par les étudiants et utilisons différents indicateurs :

- Pour mesurer l'efficacité pédagogique, nous utilisons le gain brut moyen;
- Pour mesurer l'équité pédagogique nous utilisons le coefficient de variation, avant et après le dispositif ;
- Pour mesurer l'efficience du dispositif nous utilisons le gain relatif, c'est-à-dire le rapport entre ce qui a été gagné lors du dispositif et ce qui aurait pu être gagné.

Ces indicateurs sont alors utilisés pour le pilotage de la qualité du dispositif : le gain brut, le coefficient de variation et le gain relatif permettent d'identifier des points d'amélioration. Ils sont intégrés aux bilans présentés aux instances finançant le dispositif et servent à la communication auprès des futurs étudiants (Université Paris-Saclay, 2020).

Toutefois, un certain nombre de questions doivent être discutées. Tout d'abord si la conception du dispositif est évidemment un point important, son évaluation doit également être réfléchiée dès sa conception. Par ailleurs, nous avons choisi comme données la perception de maîtrise des étudiants auto-rapportée. Il ne s'agit donc pas d'un niveau de maîtrise effectif, mais perçu. Comme le soulignent Perrault et al. (2010) « Un sentiment d'efficacité personnelle élevé comme indicateur de l'efficacité d'une formation présente des limites, il ne garantit en aucun l'existence réelle de compétence. L'inverse est également vrai ». Toutefois, il ne s'agit pas ici de considérer le niveau de maîtrise comme mesure, mais plutôt son évolution avant et après le dispositif. Notons que certains étudiants déclarent avoir le niveau maximum de maîtrise avant le dispositif, ou un niveau de perception de maîtrise inférieur après le dispositif. On peut se demander si le dispositif leur a permis de prendre conscience de leur niveau effectif de maîtrise, ce que l'on peut considérer comme un effet positif du dispositif. L'intérêt d'une auto-évaluation du niveau de maîtrise dans un dispositif d'aide aux primo-entrants, c'est à dire la fonction formative de l'auto-évaluation pourrait ainsi être discuté, mais ce n'est pas l'objet d'étude de ce travail.

Une autre interrogation porte sur le moment de l'évaluation, à chaud ou à froid ? Lorsque l'évaluation est menée à chaud, juste à la fin du dispositif (comme cela a été le cas dans les deux dispositifs analysés) il n'est pas possible de mesurer si les étudiants ont modifié leurs

pratiques de manière pérenne, ou s'ils ont pu mettre en pratique les thèmes abordés dans le dispositif à d'autres moments de leur formation. Pour mesurer un transfert des apprentissages par les étudiants une auto-évaluation à froid, en plus de celle effectuée juste à la fin du dispositif, pourrait être pertinente. Toutefois, il faudrait être vigilant sur son interprétation car d'autres éléments peuvent intervenir sur le transfert des apprentissages entre la fin du dispositif et son évaluation à froid.

Notons qu'un des obstacles pour réaliser l'évaluation d'un dispositif est celui d'avoir un échantillon représentatif des étudiants ayant bénéficié du dispositif. Le fait que les dispositifs analysés étaient obligatoires pour tous les étudiants a levé cet écueil.

5. Conclusion

L'évaluation de dispositifs pédagogiques d'aide aux primo-entrants à l'université est devenue ces dernières années une préoccupation importante des universités. Evaluer un dispositif de formation permet de l'améliorer, de le valoriser et de rendre compte aux différentes parties prenantes (étudiants, enseignants, institution l'ayant financé, ...) (Depover, et al., 2018 ; Bernard, 2011 ; Salmon et al., 2009).

Différents critères peuvent être utilisés pour évaluer un dispositif pédagogique. A partir d'une revue de littérature, nous avons retenu trois critères à évaluer pour un dispositif : l'efficacité pédagogique, l'équité pédagogique et l'efficience. Ces critères sont évalués sur la base d'indicateurs calculés à partir de la perception de maîtrise des étudiants, auto-rapportée avant et après le dispositif. Cette méthodologie a été appliquée à l'évaluation et au pilotage de deux dispositifs destinés à des primo-entrants dans deux composantes de l'Université Paris-Saclay : des ateliers en 1^{ère} année de DUT⁴ à l'IUT de Sceaux, et un module pour les étudiants inscrits en parcours « OUI-SI »⁵ à l'UFR Sciences. Les indicateurs proposés servent alors au pilotage de ces dispositifs par les formateurs, aux bilans présentés aux instances et à la communication auprès des futurs étudiants. Ainsi, le travail présenté ici s'inscrit dans le cadre des recherches évaluatives orientées vers la conception (*Design-based research*) puisqu'il s'agit de contribuer à l'amélioration de dispositifs évalués (Depover et al., 2018).

⁴ DUT : Diplôme Universitaire de Technologie

⁵ Parcours « OUI-SI » : mis en place à la rentrée 2018, ces parcours sont destinés aux bacheliers qui, en raison de leurs résultats scolaires, sont acceptés dans la filière demandée sous réserve de suivre un parcours aménagé.

Références bibliographiques

- Bernard, H. (2011). *Comment évaluer, améliorer, valoriser l'enseignement supérieur? Guide pratique à l'usage des professeurs, cadres dirigeants, responsables pédagogiques*. De Boeck Supérieur.
- Berthaud, J. (2017). L'intégration sociale étudiante : relations et effets au sein des parcours de réussite en Licence. Thèse Université Bourgogne Franche-Comté.
- Bouchard, C. et Plante, J. (2002). La qualité : mieux la définir pour mieux la mesurer, *Cahiers du Service de Pédagogie expérimentale*, Université de Liège - 11-12/2002
- Bournaud, I. et Pamphile, P. (2020). Un dispositif d'accompagnement dans la transition lycée-université (IUT) : enjeux et effets. Accepté en Juin 2020 pour publication *Revue internationale de pédagogie de l'enseignement supérieur*.
- Cosnefroy, L., Hoffmann, C., et Douady, J. (2014). L'accompagnement méthodologique. *Recherche et formation*, (77), 29-44.
- De Ketele, J. M., et Gerard, F. M. (2007). La qualité et le pilotage du système éducatif. La Qualité en éducation. Pour réfléchir à la formation de demain, 19-38.
- Depover, C., Karsenti, T. et Komis, V. (2018). La recherche évaluative. Dans T. Karsenti et L. Savoie-Zajc, *La recherche en éducation : étapes en approches* (4^e éd. p.269-287). Les Presses de l'Université de Montréal.
- Dunberry, A. et Péchard, C. (2007). L'évaluation de la formation dans l'entreprise : état de la question et perspectives. *Report, Montréal : UQAM/CIRDEP*, [www. cpmt. gov. qc. ca/publications/pdf/RECHERCHE_T3_Dunberry_Pechard. pdf](http://www.cpmt.gouv.qc.ca/publications/pdf/RECHERCHE_T3_Dunberry_Pechard.pdf).
- Figari et Remaud (2014) *Méthodologie d'évaluation en éducation et formation, ou, L'enquête évaluative*. De Boeck Supérieur.
- Gérard, F. M. (2001). L'évaluation de la qualité des systèmes de formation. *Mesure et évaluation en éducation*, 24(2-3), 53-77
- Gérard, F. M. (2003). L'évaluation de l'efficacité d'une formation. *Gestion 2000*, Vol. 20, n°3, 13-33.
- Gilibert, D., et Gillet, I. (2010). Revue des modèles en évaluation de formation : approches conceptuelles individuelles et sociales. *Pratiques psychologiques*, 16(3), 217-238.

- Michaut, C. (2003). L'efficacité des dispositifs d'aide aux étudiants dans les universités. *Recherche et formation*.
- Michaut, C. et Romainville, M. (2012). *Réussite, échec et abandon dans l'enseignement supérieur*. p. 54-68, De Boeck Supérieur.
- Monnot, A. (2014). Les pratiques d'évaluation de la formation professionnelle et leurs déterminants dans les entreprises du SBF 120. *Management Avenir*, (2), 92-111.
- Morlaix, S., et Perret, C. (2013). L'évaluation du Plan réussite en Licence: quelles actions pour quels effets ? Analyse sur les résultats des étudiants en première année universitaire. *Recherches en éducation*, Université de Nantes, pp.137-150.
- Paivandi, S. (2015). *Apprendre à l'université*. Pédagogies en développement. De Boeck supérieur.
- Perrault, B., Brassart, D. G., et Dubus, A. (2010). Le sentiment d'efficacité personnelle comme indicateur de l'efficacité d'une formation. Une application à l'évaluation de la formation des enseignants. *Actes du congrès de l'Actualité de la recherche en éducation et en formation (AREF)*, Genève.
- Perret, C. (2015). *Le Plan Réussite en Licence*. Éditions Universitaires de Dijon.
- Perret, C. et Morlaix, S. (2014). Des effets du plan réussite en licence sur la sélection universitaire en première année de licence. *Carrefours de l'éducation*, 38(2), 175-191. doi:10.3917/cdle.038.0175.
- Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in higher education. *Educational Assessment, Evaluation and Accountability*, 22(3), 215-225.
- Salmon, D., Baillet, D., Boulvain, M., Cobut, B., Coupremagne, M., Duchâteau, D., ... Houart, M. (2009). Construction d'un outil d'évaluation de la qualité des actions d'accompagnement pédagogique. Synthèse d'échanges et d'analyse de pratiques professionnelles en Communauté française de Belgique. *Revue internationale de pédagogie de l'enseignement supérieur*, 25(25-2).
- Sall, H. N., et De Ketele, J. M. (1997). L'évaluation du rendement des systèmes éducatifs: apports des concepts d'efficacité, d'efficience et d'équité. *Mesure et évaluation en éducation*, 19(3), 119-142.

Sanchez, É. et Monod-Ansaldi, R. (2015). Recherche collaborative orientée par la conception. Un paradigme méthodologique pour prendre en compte la complexité des situations d'enseignement-apprentissage. *Éducation et didactique*, 9(2), 73-94.

Saporta, G. (2006). *Probabilités, analyse des données et statistique*. Editions Technip.

Tamkin, P., Yarnall, J., et Kerrin, M. (2002). *Kirkpatrick and Beyond: A review of models of training evaluation*. Brighton, England: Institute for Employment Studies.

Université Paris-Saclay (2020), *La réussite étudiante Université Paris-Sud/ Paris-Saclay, Actions 2018-2019 et bilans après la loi ORE*. <https://www.persay.universite-paris-saclay.fr/la-reussite-etudiante-universite-paris-sud-paris-saclay-actions-2018-2019-et-bilans-apres-la-loi-ore/>