



Comment construire un processus d'évaluation en EIAH fondé sur le *Design-Based Research* ?

► **Nadine MANDRAN, Maelle PLANCHE** (Laboratoire Informatique de Grenoble LIG), **Patricia MARZIN-JANVIER** (CREAD-Université de Bretagne Occidentale), **Mathieu VERMEULEN** (IMT Lille Douai), **Aous KAROUI, Cédric d'HAM, Isabelle GIRAULT, Claire WAJEMAN, Christian HOFFMANN** (Laboratoire Informatique de Grenoble LIG)

■ **RÉSUMÉ** • Cet article propose un modèle de méthode pour construire un processus d'évaluation d'une plateforme numérique pour l'enseignement en contexte réel dans le temps. Les objectifs de cette évaluation sont : répondre aux attentes des enseignants-utilisateurs, fournir des données pour la recherche, faire évoluer techniquement la plateforme et fournir à l'institution, financeur du projet, des retours sur son utilisation. Nous illustrons cette méthode avec le projet LabNbook, concernant 157 enseignants et plus de 4500 étudiants. La méthode incluant les sept propriétés du *Design-Based Research* (DBR), nous dressons en conclusion un bilan de l'utilisation de ces dernières.

■ **MOTS-CLÉS** • Méthode, évaluation longitudinale, *Design-Based Research*

■ **ABSTRACT** • *This paper proposes a method model to build a process for evaluating a digital learning platform in a real context over time. The objectives of this evaluation are: to meet the expectations of teachers, to provide data for research, to make the platform evolve technically and to provide the institution funding the project with feedback on its use. We illustrate this method with the LabNbook project involving 157 teachers and more than 4500 students. As the method includes the seven properties of Design-Based Research (DBR), we finally report on the use of these properties.*

■ **KEYWORDS** • *method, longitudinal evaluation, Design-Based Research*

1. Introduction et contexte

Dans le cadre des Environnements Informatiques pour l'Apprentissage Humain (EIAH) et plus particulièrement de l'usage de plateformes numériques pour l'enseignement, l'évaluation de l'impact d'une telle plateforme est fondamentale pour rendre compte de sa diffusion et de son utilité. Ces évaluations permettent de recueillir des mesures et des données utiles pour les institutions (Rectorat, Université, etc.) qui en promeuvent l'utilisation, pour avoir un retour sur la diffusion et l'utilisation d'un EIAH donné. En outre, ces plateformes numériques sont parfois construites à partir de cadres théoriques et leur utilisation en contexte réel est l'occasion de produire des données pour répondre à des problématiques de recherche. Ces données fournissent aussi des informations pour faire évoluer les aspects techniques et l'ergonomie des plateformes. Les évaluations peuvent également répondre à des besoins des enseignants, en fournissant par exemple des informations sur l'activité de leurs élèves. Ces évaluations doivent être conduites dans le temps pour étudier les évolutions induites par l'utilisation et l'appropriation de la plateforme par les utilisateurs.

Le problème qui se pose est alors de savoir comment construire une évaluation longitudinale qui puisse répondre à ces différents besoins : 1) les attentes des financeurs du projet, 2) les besoins des enseignants, 3) la production de données pour la recherche et 4) l'amélioration technique de la plateforme. Ce problème est d'autant plus difficile que les personnes ayant à conduire ces évaluations sont rarement formées aux méthodes de production des données des Sciences Humaines et Sociales (p. ex., questionnaires, entretien, focus groups, etc.).

Cet article propose une méthode adaptable pour construire un processus d'évaluation longitudinale (PEL) d'une plateforme numérique conçue dans un cadre de recherche et financé par une institution. Nous présentons les différentes étapes et les outils nécessaires pour construire ce type d'évaluation. Trois questions de recherche conduisent cet article :

- 1) Quelles sont les particularités de l'évaluation des EIAH dans un contexte de recherche ?
- 2) Quelles sont les méthodes de production de données adéquates ?
- 3) Comment construire un tel processus afin de produire des données pertinentes pour l'ensemble des acteurs du projet ?

Ces questions de recherche, bien qu'elles soient généralisables à tout EIAH, sont, dans notre cas, issues du cadre théorique et d'un contexte spécifique à un projet de plateforme numérique d'apprentissage dans le supérieur. En effet, pour illustrer les étapes de cette méthode, nous l'appliquons au cas du projet LabNbook (LNB). L'objectif est la conception et la diffusion de la plateforme LNB dédiée à l'enseignement des sciences expérimentales (Biologie, Physique, Chimie, Géologie, etc.). Actuellement, cette plateforme est utilisée par 157 enseignants du supérieur et près de 4500 étudiants. Les propriétés proposées par (Wang et Hannafin, 2005) dans le cadre du *Design-Based Resarch* serviront à élaborer ce processus.

L'article est structuré en cinq sections. Tout d'abord, nous dressons un état de l'art sur l'évaluation des plateformes numérique en EIAH. Ensuite, nous nous positionnons par rapport à une évaluation en contexte de recherche ancrée dans le terrain ayant besoin de produire des données. Nous poursuivons sur la présentation de la méthode de construction du processus d'évaluation longitudinale (PEL) et de son illustration par le projet LNB. Nous ferons ensuite un retour sur les avantages et les limites de ce PEL. Nous concluons sur une série de questionnements à propos de ces évaluations en contexte de recherche et plus particulièrement sur les propriétés du *Design-Based Research* (DBR).

2. Evaluation des plateformes et DBR

Dans cette section, nous présentons un état de l'art sur les évaluations des plateformes en EIAH. Depuis l'apparition des premières plateformes de type *e-learning*, leurs méthodes d'évaluation ont évolué au fil des années. Ainsi, les approches utilisées au départ se concentraient principalement sur l'utilisabilité, en vérifiant le respect des normes théoriques (Rentroia-Bonito et Jorge, 2003) (Nielsen, 1994). Dans un deuxième temps, il y a eu plus de synergie entre les recherches sur l'utilisabilité et celles sur les EIAH en associant des tests-utilisateurs à des évaluations ergonomiques de la part d'experts (Granić et Ćukušić, 2011). Plus récemment, dans une méta-étude portant sur 30 plateformes numériques, Donnelly et al. mettent en avant une liste de critères pour identifier des plateformes efficaces (Donnelly et al., 2014), comme les fonctionnalités fournies aux enseignants, l'impact sur l'apprentissage, le développement de l'autonomie des apprenants et de pratiques d'apprentissage métacognitives, et la mise en œuvre de la collaboration. Issus de la comparaison des impacts de plusieurs plateformes sur l'apprentissage, ces critères peuvent paraître spécifiques aux plateformes numériques. Cependant ils négligent les aspects d'utilisabilité. Toutefois, ces travaux

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

portent le plus souvent sur des évaluations ponctuelles qui ne sont généralement pas conduites en situations réelles, ni sur le long terme.

Le *Design-Based Research* a été proposé par (Collective, 2003), (Collins, 1992), pour conduire des recherches en situation réelle. (Wang et Hannafin, 2005) ont étudié des recherches de terrain selon sept propriétés.

- 1) Pragmatique: il s'agit de faire progresser, d'une part, les contributions scientifiques dans l'enseignement et dans l'apprentissage et, d'autre part, la pratique des enseignants ;

- 2) Fondée sur le terrain: la recherche est conduite en « *real world where participants interact socially* » (Wang et Hannafin, 2005) ;

- 3) Interactive: les chercheurs, les enseignants, les élèves et/ou les décideurs travaillent ensemble afin de faire émerger la connaissance scientifique et la mise en œuvre de cette connaissance dans le monde réel ;

- 4) Itérative: il s'agit de différentes périodes au cours desquelles des allers-retours entre recherche et terrain vont être réalisés ;

- 5) Flexible: c'est la possibilité de faire évoluer la question de recherche par l'analyse des données expérimentales ;

- 6) Intégrative: le DBR intègre différentes méthodes de production et d'analyse des données ;

- 7) Contextuelle: il s'agit de la nécessité de faire le lien entre les résultats et le contexte dans lequel les données et les résultats ont été produits.

Plusieurs contextes doivent être étudiés; cette multiplicité de contextes permet de garantir la « généralité » des résultats et d'augmenter leur niveau de validité. A la suite de cette analyse, les auteurs discutent 9 principes rarement repris dans la littérature. Un dixième principe a été ajouté par (Mandran *et al.*, 2021) pour la mise en place d'indicateurs d'objectifs, afin de suivre et piloter les itérations.

Le DBR tend à se diffuser dans la communauté des EIAH et, bien que cette méthode soit basée sur des propriétés, elle reste sujette à de nombreuses interprétations. A ce propos, lors d'une conférence, une chercheuse a résumé la situation par « *on fait tous du DBR car l'on mixe les méthodes de production de données, on travaille in situ, on fait des itérations, on intègre tous les utilisateurs mais en dehors de ça ?* ». En effet, les propriétés bien que pertinentes restent difficiles à implémenter, car le DBR n'est pas décrit comme un processus et ne propose pas d'outils

(Herrington *et al.*, 2007), (McKenney et Reeves, 2019). Ces auteurs ont également proposé des processus et des outils pour accompagner cette méthode. Cependant, les propriétés du DBR ne sont pas clairement mises en lien avec les étapes des processus proposés.

Dans ces travaux, nous avons identifié des évaluations d'utilisabilité, d'utilité des dispositifs pour l'apprentissage et des méthodes d'évaluation itératives. En revanche, nous n'avons pas trouvé de processus d'évaluation longitudinale, c'est à dire d'organisation dans le temps de ces évaluations, ni d'évaluation qui réponde à plusieurs objectifs. Pour répondre à ce manque, nous proposons un processus d'évaluation longitudinal en lien avec les propriétés du DBR.

3. Positionnement de nos travaux

Avant de présenter la manière dont nous pouvons envisager la construction d'un processus d'évaluation dans les EIAH, nous tenons à préciser notre positionnement sur ce que signifie « évaluer dans un contexte de recherche en EIAH » et sur la production des données.

3.1. Evaluer dans un contexte de recherche en EIAH

L'un des objectifs de l'utilisation des plateformes en EIAH étant de produire des données pour la recherche, il convient de choisir une posture épistémologique pour cadrer le processus de création des connaissances scientifiques, leur domaine de validité et leur mise en œuvre dans le monde réel. Il s'agit ici de la posture du chercheur vis-à-vis de l'élaboration d'une connaissance scientifique, d'un savoir et non du cadre de transmissions des savoirs par un enseignant (De Vries et Baillé, 2006). Nous avons souvent observé une confusion entre ces deux « formes » d'épistémologie et c'est pourquoi nous tenons à préciser ici notre posture épistémologique de chercheur. « Pour un chercheur, expliciter sa posture épistémologique dans un projet de recherche, c'est alors préciser les hypothèses fondamentales sur lesquelles se fonde le processus d'élaboration et de justification des connaissances dans le projet considéré. » (Avenier et Thomas, 2015). Présenter ce cadre épistémologique du chercheur dans les publications de recherche est indispensable pour conduire une démarche scientifique rigoureuse car le chercheur se positionne sur sa manière de construire de la connaissance et sur les critères de validité de ses travaux. Par exemple, pour étudier l'impact d'une plateforme en EIAH, nous avons besoin de mener nos investigations sur du long terme dans un contexte réel auprès de plusieurs utilisateurs (étudiants et enseignants) qui vont être amenés à utiliser cette

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

plateforme et à modifier leurs manières d'enseigner ou d'apprendre. Dans ce contexte, le recours aux plans expérimentaux classiques de la statistique inférentielle n'est pas suffisant. Nous avons besoin de mobiliser à la fois des données qualitatives pour comprendre par exemple les activités des enseignants et des données quantitatives pour dénombrer. Mobiliser ces deux démarches de production de données pour répondre à des objectifs de recherche demande de choisir un cadre épistémologique qui autorise cette mixité. Cette approche pluri-méthodes (entretien, questionnaires, traces, etc.) offre au chercheur la possibilité de mobiliser toutes les sources de données pour circonscrire et étudier une question de recherche. (Balacheff et Lund, 2013) proposent ce positionnement au travers du terme « multi vocalité ».

Pour choisir un cadre épistémologique, nous proposons tout d'abord d'identifier quels connaissances et/ou outils sont produits durant ces travaux. La recherche outillée avec une plateforme numérique peut s'apparenter à une Recherche en Informatique Centrée Humain (RICH). Cette recherche se base sur l'activité humaine pour créer des dispositifs supports à cette activité. Elle produit d'une part une contribution scientifique (p. ex., modèle de l'apprenant, modèle de scénarios, etc.) et d'autre part des outils activables pour accompagner l'activité humaine (par ex., une interface, une plateforme, etc.) (Mandran, 2018). La plateforme informatique est un outil activable dans le sens où l'utilisateur va réaliser une activité avec l'outil (p. ex., enseigner, répondre à des questions pour produire un rapport, développer des scénarios d'apprentissage, produire des ressources pédagogiques, etc.).

La connaissance scientifique et l'outil sont entrelacés et interdépendants (Simon, 1996/2004), (Avenier, 2019). L'utilisation de l'outil activable dans la réalité contribue au raffinement de connaissances ou à la création de connaissances. La figure 1 représente ce cadre de recherche.

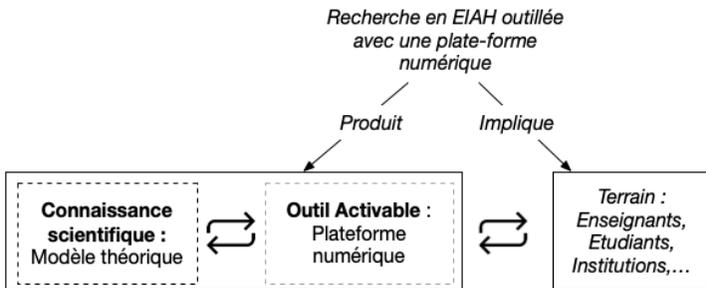


Figure 1 • Proposition d'un cadre pour une recherche en EIAH

Dans le cadre de la recherche en EIAH, il convient de décrire aussi le cadre théorique de l'apprentissage sur lequel s'appuie la construction de la plateforme et de ses fonctionnalités. Les cadres théoriques explicitent les concepts manipulés par le chercheur, permettent d'élaborer des indicateurs à mesurer sur le terrain et de faire une interprétation cohérente des données recueillies en vue de répondre aux hypothèses posées. Dans le cas de la didactique des sciences, différents cadres existent. Ils sont souvent mobilisés en fonction des problématiques (Lhoste et Orange, 2015).

La posture épistémologique du chercheur, telle que décrite précédemment, et le cadre théorique de l'apprentissage doivent être choisis avant d'initier tout processus d'évaluation, afin d'obtenir des données pour répondre à des questions de recherche.

3.2. Données et outils de mesures

Dans cette section, nous indiquons notre positionnement sur certains concepts liés aux données. Pour répondre à une problématique de recherche, des données doivent être recueillies ou produites. Selon l'OCDE : «Les données de la recherche sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche». Le concept d'*enregistrement factuel* traduit la nécessité de garder une trace d'une observation réalisée sur le terrain pour valider des résultats scientifiques. Mais cette définition ne précise rien sur la nature des données. D'autres définitions sur les données existent, mais elles restent relativement larges et sont dépendantes des domaines scientifiques.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Nous avons donc choisi un positionnement vis-à-vis de ces concepts de données et de mesures. Nous faisons la distinction entre des données qui existent avant la recherche et celles qui seront produites par la recherche. Par exemple, un support de cours (comme un manuel scolaire) est une « **donnée existante** » hors du contexte de la recherche, que le chercheur peut mobiliser pour répondre à ses questions. Le terme de « **donnée mesurée** » sera utilisé pour désigner les données construites par et pour le chercheur sur la base de travaux théoriques (p. ex., une note d'utilisabilité). Ces deux notions font la distinction entre ce qui est créé dans une démarche scientifique et ce qui est issu du terrain indépendamment d'une question de recherche.

Pour produire ces données mesurées, à l'instar des sciences expérimentales, le chercheur conçoit **un outil de mesures** pour recueillir des données. Il peut s'agir d'un questionnaire ou d'un guide d'entretien. Lors de la passation d'entretiens, les guides d'entretiens garantissent la reproductibilité de la méthode. Il en est de même pour les questionnaires lors d'enquêtes. Un autre chercheur peut utiliser ces mêmes outils pour la même question de recherche.

Pour la production des données mesurées, nous faisons également deux autres distinctions. D'une part, nous distinguons la démarche **qualitative** de la démarche **quantitative**. La démarche qualitative (Paillé et Mucchielli, 2011) permet de comprendre et d'identifier des phénomènes. Le nombre d'utilisateurs consultés est généralement faible (entre 6 et 20 personnes). En revanche, la variété des profils est importante pour observer un maximum de cas. Les méthodes de production des données sont, par exemple, des observations *in-situ*, des entretiens ou des *focus-groups*. Les méthodes d'analyse des données mesurées sont alors du codage de corpus, de l'analyse thématique ou de l'analyse textuelle. La démarche quantitative (Creswell, 2013) permet de quantifier des faits ou des avis. La sélection des individus repose soit sur une méthode d'échantillonnage (i.e. quotas, aléatoire) soit sur un recensement. Dans ce cas, l'ensemble des personnes est consulté. Les données mesurées sont, par exemple, des traces d'activités ou des réponses à des questionnaires. Les méthodes d'analyse des données sont des analyses statistiques et des visualisations.

D'autre part, pour caractériser la nature des mesures, nous distinguons la nature **factuelle** de la nature **déclarative** des données mesurées. Les

mesures factuelles résultent de l’observation de faits ou d’activités. Elles sont mesurées par des traces d’activité sur la plateforme ou des captures vidéo de séances d’enseignement. Les mesures déclaratives reposent sur l’expression des enseignants et des étudiants sur leurs expériences et leurs perceptions. Elles sont recueillies par des questions dans des questionnaires ou lors d’entretiens.

4. Comment construire un Processus d’évaluation longitudinale (PEL) ?

La méthode que nous proposons a pour objectif de guider des chercheurs dans la construction d’un PEL afin de répondre aux quatre objectifs décrits précédemment (institution, enseignement, technique et recherche). Pour la formaliser, nous utilisons le modèle de méthode THEDRE (*Traceable Human Experiment Design REsearch*) (Mandran, 2018), (Mandran et Dupuy-Chessa, 2017), qui a été conçu pour apporter des outils de traçage à des travaux conduits dans le domaine de l’informatique impliquant l’humain. Il est fondé sur la théorie du constructivisme pragmatique, de la démarche qualité et de la démarche centrée utilisateurs. Il permet de dresser une liste de tâches pour conduire un processus de recherche et de lui adjoindre des outils facilitateurs (p. ex., un guide pour rédiger un protocole de production des données). Dans le cadre de cet article, nous présentons notre méthode de construction du PEL sous la forme d’une liste de tâches structurée en six temps (voir Fig. 2).

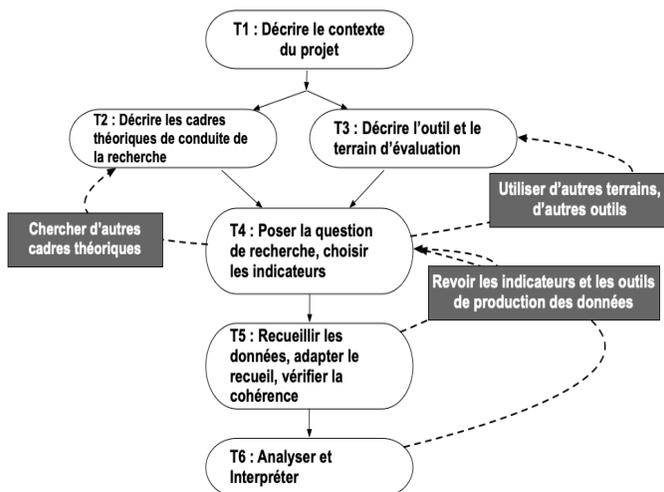


Figure 2 · Description du processus de construction d'un dispositif d'évaluation d'une plateforme numérique en EIAH

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Avant d'exemplifier ce processus par le cas de la plateforme LabNbook, nous expliquons chacune des tâches pour présenter et expliciter le processus dans sa globalité.

TEMPS 1 : « Décrire le contexte du projet »

- T1.1: Décrire l'organisation du projet et pour qui l'évaluation doit être réalisée. Cette étape permet de circonscrire les objectifs du projet, les personnes impliquées et les contraintes. Elle permet aussi d'identifier à quels acteurs du projet l'évaluation fournira des informations et dans quelle finalité.

- T1.2: Se mettre en accord avec les règles éthiques et de déontologie ainsi qu'avec la réglementation de production des données.

TEMPS 2 : « Décrire les cadres théoriques de conduite de la recherche »

- T2.1: Choisir une posture épistémologique pour construire la connaissance scientifique (voir 3.1).

- T2.2: Choisir une méthode de conduite de la recherche. Identifier des principes ou des processus à suivre pour conduire une recherche. (par ex., *Design-Based Research*).

- T2.3: Décrire dans quel cadre théorique de l'apprentissage l'outil numérique a été élaboré et dans lequel il sera évalué. Il s'agit de décrire le cadre épistémologique de transmission des savoirs, ainsi que d'autres cadres théoriques de référence. Ces cadres permettent d'élaborer des questions de recherche et d'avoir des moyens de produire et d'analyser les données.

TEMPS 3 : « Décrire l'outil et le terrain d'évaluation »

- T3.1: Décrire l'outil numérique et ses fonctionnalités. Cette description permet d'identifier les fonctionnalités à évaluer.

- T3.2: Décrire comment le terrain sera pris en compte, avec quels utilisateurs, et les contraintes inhérentes à ce terrain. Il s'agit d'identifier les utilisateurs à impliquer et leurs contraintes, de manière à anticiper les problèmes de recrutement ou de calendrier et trouver des alternatives. Les biais expérimentaux sont décrits dans cette étape.

TEMPS 4 : « Poser les questions de recherche, choisir les indicateurs »

- T4.1: Poser les questions de recherche. En lien avec les cadres théoriques de l'apprentissage mobilisés et l'outil proposé, il s'agit de rédiger la problématique, les questions de recherche auxquelles l'évaluation apportera des réponses.

- T4.2: Choisir les indicateurs, construire une liste de mesures à prendre sur le terrain en lien avec ces indicateurs, les répertorier par objectifs.

- T4.3: Choisir les types de données à produire en lien avec la posture épistémologique, la méthode de conduite de la recherche (p. ex., *Design Based Research*) et la démarche de production des données (p. ex., qualitatif, quantitatif, mixte, etc.).

- T4.4: Identifier les données existantes sur le terrain, à recueillir auprès des acteurs.

- T4.5: Construire les outils de mesures associés, les tester et les améliorer. Il s'agit de disposer des outils suffisamment décrits et évalués avant la production des données sur le terrain. Ce travail est le plus souvent itératif. Les biais expérimentaux sont complétés dans cette étape.

- T4.6: Schématiser le processus. Il s'agit d'avoir une vue globale des étapes, des objectifs et des méthodes de production des données. Représenter graphiquement le processus d'évaluation pour le suivre, le partager, le diffuser.

TEMPS 5: «Recueillir les données, adapter le recueil et vérifier la cohérence»

- T5.1: Recueillir les données existantes et mesurées sur le terrain.

- T5.2: Faire un bilan sur les données recueillies et un diagnostic sur les outils de mesures, les adapter, les améliorer, les supprimer. Il s'agit de vérifier que les données recueillies sont cohérentes avec les objectifs fixés et que les outils de mesures sont adaptés à la situation.

A la suite de ce temps de recueil, une dernière étape consiste à analyser et interpréter les données. Ce sixième temps ne sera pas exemplifié dans cet article. Cependant, à titre indicatif, nous proposons trois tâches.

TEMPS 6: « Analyser et interpréter »

- T6.1: Réaliser une analyse descriptive des données recueillies. L'objectif est d'avoir un premier état des lieux des données recueillies. Cette analyse permet d'avoir un premier niveau d'interprétation de ce qui a été mesuré. Elle offre déjà la possibilité de répondre à la question de recherche, voire de la compléter, la raffiner.

- T6.2: Choisir les méthodes d'analyse en lien avec le type de données ou de mesures recueillies. Le choix de ces méthodes d'analyse permet d'élaborer un plan de traitement des données et de le documenter.

- T6.3: Analyser les données et mesures, interpréter les résultats.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Globalement les six temps doivent se conduire dans l'ordre car ils dépendent les uns des autres. En effet, il est nécessaire d'explicitement le contexte du projet et les cadres théoriques pour choisir les indicateurs répondant aux objectifs. Cependant des allers-retours entre les différents temps sont possibles comme exposé sur la Figure 2. Poser la question de recherche et le choix des indicateurs peut demander de revenir interroger les cadres théoriques pour raffiner la question et choisir d'autres indicateurs. Le recueil de données peut demander de revoir les outils de production utilisés. Un problème de terrain peut demander de choisir un autre terrain d'évaluation ou une autre plateforme numérique. Quant aux tâches à l'intérieur des temps, elles peuvent se conduire en parallèle et de manière itérative. Par exemple, la création d'un questionnaire va réinterroger les indicateurs à produire.

5. Application au Projet IDEX – LabNbook (LNB)

Cette section présente une illustration de la construction d'un processus d'évaluation au travers du projet LNB.

5.1. TEMPS 1 : Décrire le contexte du projet

Le processus d'évaluation dont il est question ici a été construit dans un contexte de projet financé par l'IDEX formation de l'Université Grenoble Alpes autour de la plateforme LNB. L'objectif pédagogique de LNB est de former les étudiants à la démarche expérimentale en sciences. Cet outil a pour finalité d'accompagner des enseignants des sciences expérimentales dans la conception d'enseignements centrés sur l'étudiant et de permettre aux étudiants de produire de façon collaborative des écrits scientifiques (p. ex., cahiers de laboratoire, rapports, etc.).

Le projet vise à promouvoir les transformations pédagogiques sur le site universitaire de l'Université Grenoble Alpes (UGA) en particulier vers des pédagogies dites « actives », à l'aide de LNB. L'objectif principal est d'étudier les pratiques des enseignants, les usages et l'appropriation de la plateforme par des enseignants et des étudiants en sciences expérimentales, au fur et à mesure de l'utilisation de LNB. Il doit répondre à quatre acteurs : les financeurs du projet, les enseignants utilisateurs de la plateforme, les chercheurs en didactiques des sciences expérimentales et les développeurs.

L'équipe projet est composée de cinq enseignants-chercheurs qui ont conçu LNB. La programmation et le développement de LNB repose sur l'un des enseignants-chercheurs et un informaticien. Deux méthodologues contribuent au projet pour la conception et la mise en œuvre du processus d'évaluation. Les objectifs sont 1) de fournir des retours aux financeurs sur l'utilisation et l'impact de la plateforme, 2) de répondre aux attentes et besoins des enseignants, 3) de répondre à des questions de recherche et 4) de fournir des informations aux développeurs pour l'amélioration de la plateforme. La déclaration de conformité au Règlement Général de la Protection des Données (RGPD) et les conditions générales d'utilisation (CGU) ont été rédigées en collaboration avec le responsable de la protection des données de l'UGA.

5.2. TEMPS 2 : Décrire les cadres théoriques de conduite de la recherche

5.2.1. Choisir une posture épistémologique

Notre situation de recherche sur LNB est dépendante de la réalité du terrain, elle ne peut être abordée par des méthodes expérimentales classiques. Par exemple, la présence d'un groupe contrôle versus un groupe expérimental n'est pas envisageable. En effet, si l'utilisation de la plateforme est un avantage pour l'apprentissage, nous ne pouvons pas concevoir un plan expérimental dans lequel certains étudiants n'auraient pas accès à cette plateforme sous prétexte qu'ils sont des « étudiants-test ». Cependant en tant que chercheur, il faut garantir la scientificité des résultats pour ces situations de recherche largement répandues dans le contexte de la recherche en éducation.

Pour aborder ces recherches, Avenier et Thomas (Avenier et Thomas, 2015) proposent d'utiliser le constructivisme pragmatique comme posture du chercheur vis à vis du terrain. Cette posture se définit selon cinq hypothèses principales :

- 1) « *l'expérience humaine est observable dans le système auquel elle appartient* ». Dans le cas de LNB, nous pouvons impliquer et observer les étudiants et les enseignants.

- 2) « *le monde observé concerne l'humain dans un contexte donné* ». Dans notre cas, nous allons observer des étudiants et des enseignants dans une salle de classe ou hors de la classe.

- 3) « *Ce contexte peut varier ; les actions et les décisions de l'humain peuvent évoluer* ». Avec LNB, les enseignants peuvent choisir de modifier les ressources fournies aux étudiants au cours de l'enseignement.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

- 4) « *La contribution prend la forme de modèles génériques activables* ». Dans notre cas, notre recherche donnera lieu au raffinement d'un modèle théorique de l'apprentissage (voir 5.2.3)

- 5) « *La contribution scientifique et les questions de recherche sont flexibles et vont évoluer en fonction du terrain* ».

Dans le constructivisme pragmatique, le but de la construction scientifique est « *d'élaborer des modèles intelligibles de l'expérience humaine, offrant des repères adaptés et viables* », « *La production de connaissances scientifiques contribue à la génération ou au raffinement de connaissances* ». Dans notre cas, nous souhaitons contribuer à l'amélioration d'un modèle d'enseignement des sciences expérimentales selon une pédagogie active outillée par LNB. Concernant la valeur et la validité des résultats, nous nous positionnons sur les propositions du constructivisme pragmatique (Avenier et Thomas, 2015). Ainsi, lors de notre évaluation, nous vérifierons que l'outil est en adéquation avec les activités des enseignants et des étudiants et nous mobiliserons à la fois des méthodes de production de données qualitatives et quantitatives pour cerner au mieux le terrain.

5.2.2. Choisir une méthode de recherche

Les hypothèses de construction de la connaissance scientifique et les critères de valeur et validité étant choisis, il convient de raffiner cette posture dans un cadre dédié à la recherche en éducation outillée par des plateformes numériques. Nous avons choisi le *Design-Based Research* et nous en utiliserons les sept propriétés pour qualifier notre processus d'évaluation lors de la description du terrain d'évaluation.

5.2.3. Décrire le cadre théorique de l'apprentissage

Le propos de l'article n'étant pas les résultats de cette évaluation nous décrivons brièvement les cadres théoriques.

Le cadre théorique de l'apprentissage de LNB repose sur le socio-constructivisme (Vygotsky, 1978). Deux cadres théoriques sous-tendent nos recherches, d'une part le cadre de la théorie des deux mondes (Tiberghien *et al.*, 2009) et d'autre part la « pédagogie active » (Hannafin *et al.*, 1999), auquel nous associons le cadre du *Student-Centered Learning* (SCL) (Freeman *et al.*, 2014). Le cadre de la théorie des deux mondes explique l'apprentissage des sciences expérimentales par la manipulation

de relations entre le monde des objets et événements et le monde des théories et modèles ; il repose principalement sur deux sources théoriques générales : une approche anthropologique des savoirs et une approche épistémologique de l'enseignant considérant la modélisation du monde physique comme une activité centrale. Ces deux sources complètent l'approche socio-constructiviste. L'approche par pédagogie active est définie comme « *a general design framework to describe environments that support personal sense making via problems contexts enriched with technology tools, resources and scaffolding* » (Hannafin *et al.* 1999). Ces environnements ont les caractéristiques suivantes : ils visent l'autonomie et l'indépendance de l'apprenant en lui déléguant en partie la responsabilité de son parcours d'apprentissage (Lathika, 2016). En s'appuyant sur ces principes, (Land et Oliver, 2012) et (Lee et Hannafin, 2016) ont défini les caractéristiques de ces environnements avec cinq critères :

- Critère 1 - Une situation-problème ou un projet authentique et concret, au regard des apprentissages visés, sont proposés aux élèves, à partir desquels plusieurs solutions sont possibles (Hannafin *et al.*, 1999) ;

- Critère 2 - Les apprenants doivent pouvoir utiliser leurs connaissances initiales pour entrer dans la tâche ;

- Critère 3 - Les ressources peuvent être conceptuelles (les notions en lien avec le problème posé), métacognitives (comment expliciter, réfléchir, planifier ou contrôler son travail), procédurales (comment utiliser les fonctionnalités), et stratégiques (comment appréhender le problème) ;

- Critère 4 - L'apprenant doit également pouvoir faire des choix et être guidé dans ces choix ;

- Critère 5 - L'apprenant doit pouvoir évaluer ce qu'il sait et ce qu'il a besoin de savoir pour atteindre le but visé (Hannafin et Hannafin, 2010).

Ces éléments de cadrage aideront à la construction des questions de recherche et des indicateurs à mesurer sur le terrain lors de l'utilisation de LNB. Ils contribueront au guidage de l'analyse des données (p. ex., construction d'une grille de codage pour l'analyse).

5.3. TEMPS 3 : Décrire l'outil numérique et le terrain

Ce temps est consacré à la description de l'outil numérique utilisé en contexte réel et sur le terrain d'évaluation mobilisé.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

5.3.1. Décrire l'outil numérique

En lien avec ces cadres théoriques, qui mettent la focale sur l'apprenant, ses connaissances initiales, et sur l'utilisation de démarches expérimentales pour apprendre, un ensemble de fonctionnalités et d'outils sont proposés sur LNB. La plateforme permet aux étudiants de produire et de partager des cahiers de laboratoire ou des comptes rendus expérimentaux. LNB leur fournit des outils de production et de communication, pour produire des écrits d'une démarche expérimentale. Sur LNB, un même espace de travail peut être partagé par une équipe d'étudiants ; les documents écrits sont le résultat d'un travail collaboratif. Les enseignants construisent une mission (i.e. la situation d'enseignement) qui correspond à l'espace de travail des étudiants. Il est structuré en un ensemble de parties qui sont complétées en mobilisant quatre outils de production, appelés *labdocs* (traitement de textes, dessins, traitement de données, écriture de protocoles expérimentaux) (critère 3). Au moment de configurer la mission, l'enseignant choisit ce qu'il souhaite laisser à la charge de l'étudiant et ce qu'il souhaite guider selon ses priorités pédagogiques. LNB offre diverses possibilités de guidage dans une mission. L'enseignant peut structurer le rapport en choisissant le nombre de parties et les outils mobilisables par l'étudiant, et en proposant des *labdocs* modifiables par les étudiants (ils peuvent être vides et n'indiquer que le titre, ou bien présenter un contenu). Sur LNB, il est possible d'explicitier les objectifs d'apprentissage ou les compétences visées dans différents espaces de la plateforme, de mettre à disposition des étudiants des documents et des consignes de travail (critère 3). Les enseignants peuvent suivre le travail des étudiants grâce à un tableau de bord et faire des rétroactions sur les productions des étudiants en les annotant, en fonction des réussites, des erreurs et des difficultés identifiées (critères 2, 4 et 5).

5.3.2. Décrire le terrain et les utilisateurs

Le choix du DBR, comme méthode de conduite de la recherche, nous incite à décrire notre terrain selon les sept propriétés du DBR.

Plusieurs chercheurs sont mobilisés autour de la recherche sur LNB. Ils sont eux-mêmes enseignants et utilisent la plateforme avec leurs étudiants. Un d'entre eux est le concepteur de LNB et il participe au développement. (*Propriétés du DBR : interactive*).

Plusieurs types d'utilisateurs sont mobilisés pour évaluer l'impact de LNB (*Propriétés du DBR : interactive*) :

- Les concepteurs des missions. Il s'agit de responsables d'unité d'enseignement (UE) qui impulsent l'utilisation de LNB et conçoivent les missions qui seront utilisées par les enseignants de l'équipe pédagogique. Nous étudions leurs pratiques et motivations, leur niveau d'appropriation et de satisfaction. Ils sont sollicités pour contribuer à l'amélioration de la plateforme lors de séminaires. Ils font progresser LNB et les questions de recherche (*Propriétés du DBR : interactive*).

- Les tuteurs. Il s'agit de l'ensemble des enseignants appartenant aux équipes pédagogiques. Pour chaque enseignement, c'est le concepteur qui leur propose d'utiliser LNB et qui fournit la mission. Il arrive cependant que certains tuteurs adaptent les contenus selon leurs objectifs pédagogiques personnels. Nous quantifions leurs pratiques et motivations, leur niveau d'appropriation et de satisfaction.

- Les étudiants. Nous mesurons leur niveau de satisfaction après l'utilisation de LNB et l'apport du *Student Centered Learning* de LNB.

La plateforme est aujourd'hui déployée auprès de 51 unités d'enseignement (UE). Elle a été utilisée par 157 enseignants et 4500 étudiants depuis 2017. Ce contexte est riche et permet de recueillir un nombre important de données. (*Propriétés du DBR : pragmatique, fondée sur le terrain, interactive*).

LNB est utilisé plusieurs fois par les enseignants avec des étudiants différents, dans des disciplines différentes. Cette diffusion permet d'étudier la généricité de LNB (*Propriétés du DBR : contextuelle*).

L'implication et les rôles de ces utilisateurs étant très différents, nous avons besoin de mobiliser à la fois des méthodes qualitatives pour comprendre finement les motivations et les pratiques des concepteurs et des tuteurs et des méthodes plus quantitatives pour analyser le niveau de satisfaction (*Propriétés du DBR : intégrative*).

Le PEL doit mesurer l'impact de la plateforme au fil du temps. Les données doivent donc être recueillies avec une même cohorte d'utilisateurs, à différents moments de leur utilisation : avant, pendant et après utilisation. Ces mesures sont répétées pour chaque semestre. Nous devons pour cela disposer d'outils de mesure qui seront réutilisés au cours du temps. Les résultats obtenus à l'issue de ces phases de recueil de données vont faire évoluer la plateforme, les questions de recherche voire les outils de mesures eux-mêmes (*Propriétés du DBR : itérative et flexible*).

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Notre terrain d'évaluation répond aux sept propriétés du DBR. Nous devons maintenant poser les questions de recherche et élaborer des indicateurs qui seront mesurés sur le terrain.

5.4. TEMPS 4 : Poser les questions de recherche, choisir les indicateurs

Dans ce temps nous élaborons les questions de recherche et identifions les données nécessaires pour y répondre.

5.4.1. Poser les questions de recherche

Cinq questions de recherche sont construites à l'aide des cadres théoriques :

- Quels sont les facteurs qui déterminent l'acceptation de la plateforme LNB par les enseignants ?
- Quels supports LNB apporte-elle aux enseignants et apprenants engagés dans des enseignements « centrés-étudiant » ?
- Quels sont les types de tâches que les enseignants donnent à réaliser aux étudiants ?
- Quelle forme prend la collaboration entre étudiants pour l'écriture de documents scientifiques sur LNB ?
- Quelles modifications la plateforme LNB induit-elle dans le travail des enseignants et des étudiants ?

5.4.2. Choisir les indicateurs

A partir de ces questions de recherche et des critères définis dans ces cadres (comme la présence d'une activité de résolution de problème déléguée aux élèves), nous construisons les indicateurs à mesurer sur le terrain. L'évolution de ces différents indicateurs nous permettra de répondre aux questions de recherche sur les changements dans les activités.

Nous avons tiré de la littérature plusieurs indicateurs pour répondre aux autres objectifs : « enseignement », « technique » et « institutionnel ». Des indicateurs liés à l'utilisation, l'utilisabilité et la satisfaction ont été choisis et adaptés à notre contexte. Pour l'utilisabilité, nous utilisons l'échelle SUS (Brooke, 1996); pour la mesure de satisfaction, nous mobilisons des travaux en sciences de gestion (Ray, 2001). Les indicateurs d'utilisation de la plateforme sont élaborés par les porteurs du projet pour suivre le projet et tester la plateforme.

Le tableau 1 présente des indicateurs utilisés pour évaluer LNB. La troisième colonne du tableau donne des exemples de données mesurées.

Tableau 1 • Indicateurs et finalités

<i>Indicateurs</i>	<i>Finalité</i>	<i>Exemples de mesures</i>
Utilisation de la plateforme	Mesurer l'utilisation de la plateforme et sa diffusion dans l'institution.	Nombre d'utilisateurs par an, nombre d'unités d'enseignement, etc.
Satisfaction	Mesurer la satisfaction perçue par les étudiants et les enseignants.	Note de satisfaction globale, note de satisfaction sur les ressources fournies, etc.
Utilisabilité	Mesurer si LNB est utilisable par les enseignants et les étudiants. Connaître les difficultés dans l'utilisation.	Echelle de Brooke
Activités	Mesurer les actions réalisées avec LNB par les enseignants et les étudiants.	Liste des actions réalisées par les enseignants, les étudiants
Indicateurs du SCL (références)	Mesurer si les fonctionnalités de type SCL sont perçues par les étudiants et les enseignants.	Fondé sur l'article de Freeman et al. , 2014

5.4.3. Choisir les types de données et les méthodes de production des données

Le cadre épistémologique et la méthode de conduite de la recherche que nous avons choisis nous autorisent à utiliser plusieurs types de méthodes de production de mesures pour investiguer le terrain (*Propriété du DBR : intégrative*). Cette section présente les différentes méthodes retenues en fonction du contexte de notre terrain d'évaluation. Nous utilisons des approches qualitatives pour comprendre et des méthodes quantitatives pour dénombrer.

- Les **entretiens semi-directifs** visent à recueillir le plus précisément possible l'avis des responsables d'UE impliqués dans le projet. Ils ont été réalisés avant utilisation de la plateforme afin d'identifier leurs pratiques et leurs motivations à utiliser LNB. En fin de projet, ils sont de nouveau interrogés sur ces mêmes indicateurs. Un guide d'entretien a été construit pour conduire ces entretiens. Les données recueillies sont de type qualitatif et déclaratif et elles font l'objet d'une analyse thématique.

Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER, Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle GIRAULT, Claire WAJEMAN, Christian HOFFMANN

- Des **cahiers de bord** (Bernhaupt *et al.*, 2008) permettent de suivre les besoins d'aide des enseignants. Ils sont remplis par l'équipe conceptrice de LNB tout au long de l'utilisation. Il s'agit de données qualitatives déclaratives.

- Par les **traces numériques** des enseignants et des étudiants, nous pouvons explorer l'utilisation réelle de LNB et l'évolution de l'utilisation. Ce sont des mesures de type factuel. Pour les enseignants, nous traçons leurs activités sur les missions, sur les productions des étudiants et sur la gestion des étudiants. Pour les étudiants, nous recueillons leurs activités sur les rapports, les ressources, et les échanges.

- **L'observation des « missions »**. Il s'agit de données factuelles produites par l'enseignant sans intervention du chercheur. L'organisation et la modification des missions proposées par les enseignants nous renseignent sur le type de pédagogie mobilisée. Ainsi, pour chaque situation pédagogique proposée, nous regardons la nature du travail demandé aux étudiants et sa place dans le processus pédagogique : préparer le TP, refaire le travail, diffuser ses résultats, etc. À la fin de la seconde année, ces missions sont de nouveau analysées afin d'étudier les transformations pédagogiques.

- Nous utilisons également des **débriefings** pour identifier rapidement des points de blocage dans l'utilisation de LNB.

- Un **questionnaire-baromètre (QB)** est un questionnaire soumis de façon répétée et à intervalles réguliers à une cohorte d'utilisateurs. Dans le cas de LNB, il y a trois vagues de passation : avant l'utilisation et à la fin de chaque année d'enseignement. Ces questionnaires sont administrés auprès des enseignants concepteurs et des enseignants tuteurs, afin de mesurer l'évolution de leurs activités et de leur satisfaction. Les mesures recueillies sont de type déclaratif ; elles font l'objet de traitement quantitatif.

- Pour les étudiants, le **questionnaire (QS)** n'est pas « barométré », car d'un semestre à l'autre, ce ne sont pas les mêmes cohortes qui sont interrogées.

5.4.4. Identifier les données existantes

Les données existantes sont les supports de cours produits par les enseignants. Ils sont recueillis lors des entretiens.

5.4.5. Construire les outils de mesures

Nous présentons à titre d'exemple l'élaboration du questionnaire baromètre qui contient des indicateurs théoriques (SCL) et des indicateurs de terrain (Satisfaction). Les données produites sont quantitatives et déclaratives. Pour les indicateurs du SCL, nous utilisons les caractéristiques fournies par les auteurs (Freeman et al, 2014) et nous les déclinons en questions. Le tableau 2 présente un exemple de ce travail. Pour les réponses, nous utilisons une échelle d'accord de Likert à quatre degrés (*Tout à fait, Plutôt, Plutôt pas, Pas du tout d'accord*) et la modalité « non réponse ».

Tableau 2 • Trois Indicateurs du SCL et questions associées

<i>Trois indicateurs du SCL</i>	<i>Questions associées</i>
Les modalités de résolution du problème déléguées aux étudiants	« Pour cette formation, je laisse les étudiants choisir librement leur méthode pour résoudre des problèmes complexes »
Les étudiants sont autonomes dans leur travail	« LNB aide les étudiants à organiser leur travail » « LNB aide les étudiants à structurer leurs rapports »
Une organisation temporelle flexible du travail (adaptation au rythme de l'élève, possibilité de travail à distance, ...)	« Pour cette formation, je laisse la liberté aux étudiants d'organiser leur temps pendant la séance » « Pour cette formation, je demande aux étudiants du travail en dehors des séances » « LNB facilite le travail des étudiants entre les séances »

La mesure de la satisfaction s'effectue par une question qui mesure la satisfaction globale et par plusieurs questions qui mesurent la satisfaction sur des éléments plus fins (Ray, 2001), (Parahoo *et al.*, 2016). Ainsi, dans le cas de LNB, nous avons mesuré la satisfaction globale sur une échelle allant de 0 à 10. Nous avons posé des questions de satisfaction sur les fonctionnalités de LNB, sur l'écriture des rapports avec LNB, sur l'utilisation de LNB avec les enseignants, pour le travail à la maison, pour obtenir des retours de la part des enseignants, pour travailler avec les étudiants du groupe, pour partager des ressources. Ainsi, avec ces sept questions, nous pouvons étudier l'impact de ces éléments sur la satisfaction globale. Concrètement, nous saurons si le partage des ressources entre les étudiants (une des fonctionnalités de LNB) a un impact fort ou non sur la satisfaction globale.

Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER, Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle GIRAULT, Claire WAJEMAN, Christian HOFFMANN

5.4.6. Schématiser le processus d'évaluation

La représentation graphique du processus d'évaluation (Figure 3) offre une vision globale et une synthèse de ce travail de terrain. Au cours de ce projet, nous avons utilisé plusieurs représentations graphiques en fonction des canaux de médiation (institution, publications, web, etc.).

	Avant utilisation	Année 1	Année 2	Année 3
Enseignants concepteurs	QB0 Entretien T0	QB0 ou QB1 Missions 1	QB1 ou QB2 Missions 1 et 2 Entretien T1	QB2 Missions 2 Entretien T1
Enseignant utilisateurs	QB0	QB0 ou QB1 Débriefing 1	QB1 ou QB2 Débriefing 1 ou 2	QB2 Débriefing 2
Etudiants		Traces 1	QS1 Traces 2	QS2 Traces 3

Figure 3 • Schématisation du PEL

6. Déroulement du PEL dans le projet LNB et bilan

Dans cette section, nous présentons le déroulement sur le terrain et dressons le bilan de cette étude.

6.1. Déroulement du PEL dans le projet LNB

Le processus d'évaluation longitudinale a été déployé sur une période de trois ans, de 2017 à 2019. Au départ du projet, vingt-trois unités d'enseignement étaient impliquées. Le PEL a débuté par la passation des entretiens avec les responsables d'UE avant qu'ils élaborent les situations pédagogiques proposées à leurs étudiants sur la plateforme. À partir de septembre 2017, les questionnaires baromètres ont été soumis à l'ensemble des enseignants-utilisateurs et ce, jusqu'à décembre 2019. Les entretiens de fin d'utilisation avec les responsables d'UE ont été conduits de manière échelonnée entre les mois de mai 2019 et le mois de février 2020. Le recueil des traces numériques et des données issues des débriefings s'est opéré de façon continue tout au long du processus. Plusieurs publications de recherche ont été réalisées sur la base de ces données (Planche *et al.*, 2019), (Planche *et al.*, 2021), (Hoffmann *et al.*, 2019), (Hoffmann *et al.*, 2021).

6.2. Bilan sur les outils des mesures

Les entretiens ont principalement servi l'objectif « recherche » du PEL et, plus occasionnellement, l'objectif « technique » et « enseignement ». Nous avons réalisé 22 entretiens auprès des enseignants-concepteurs avant l'utilisation de LNB et 19 après (parmi eux 14 ont été rencontrés avant et après l'utilisation de LNB). Nous en avons conduits auprès de 5 enseignants-tuteurs impliqués dans la conception, afin d'avoir un regard complémentaire. Certains concepteurs n'ont pas souhaité continuer ; leur choix d'abandonner nous a conduits à les interroger à nouveau pour comprendre cette décision. L'analyse thématique des entretiens a éclairé le choix, les motivations et les changements des pratiques des responsables d'UE. La difficulté inhérente à ces données a été le recueil et l'analyse qui est par nature chronophage.

Les questionnaires baromètres ont essentiellement servi les objectifs « enseignement » et « recherche » du PEL. En janvier 2020, sur les trois vagues, 159 questionnaires avaient été complétés par 66 enseignants-utilisateurs (soit 42% de l'ensemble des enseignants-utilisateurs). Après nettoyage de la base et suppression des doublons, 151 questionnaires se sont avérés exploitables. Parmi les 66 enseignants ayant répondu au moins une fois au questionnaire-baromètre, 32 ont répondu aux deux premières vagues (QB0 et QB1) - soit un taux de réponse de 48,5% - et, parmi eux, 26 se sont prêtés à l'exercice pour les trois vagues de passation (QB0, QB1 et QB2) - soit un taux de réponse de 39,4%. Bien que faible sur la totalité du PEL, les réponses à ces questions vague par vague montrent comment s'est passée l'évolution des utilisations de LNB. Si ces données chiffrées sont des indicateurs sur l'évolution de l'utilisation de LNB, il convient de les étayer par des données qualitatives comme les entretiens ou des données existantes produites par les enseignants, comme les missions. Le questionnaire baromètre qui nous apparaissait comme un outil pertinent pour mesurer de manière quantitative les évolutions s'est avéré lourd et peu satisfaisant. Les trois vagues de questionnaires ont lassé les enseignants et le taux de participation n'a jamais été très élevé et a diminué à chacune des vagues. En outre, lors de l'analyse des réponses aux questionnaires, des questions nous sont apparues mal formulées (p. ex., « J'ai travaillé de façon collaborative », le terme collaboratif a sans doute été compris différemment selon les unités d'enseignement.

L'étude des missions sert principalement l'objectif « recherche » du PEL. 144 missions implémentées sur LNB ont été analysées dans le cadre

Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER, Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle GIRAULT, Claire WAJEMAN, Christian HOFFMANN

du PEL. Ces missions renseignent sur le degré de liberté laissé aux étudiants dans la conduite de leur travail par les enseignants. La structuration de ces missions et leur évolution renseignent de manière factuelle sur l'activité de l'enseignant. Une grille de codage des missions a été élaborée sur la base des critères du SCL. L'analyse de ces données a été enrichie par l'analyse thématique des entretiens.

Initialement, des carnets de bord avaient été élaborés pour suivre les interactions entre les enseignants et l'équipe conceptrice de LNB tout au long du projet. Ils devaient être complétés par l'ensemble de l'équipe conceptrice, à mesure que les demandes d'aide se présentaient. Le recueil a finalement été abandonné, en raison notamment de l'aspect chronophage et de la rigueur presque quotidienne exigée par la tâche, compte-tenu des autres charges pesant déjà sur les membres de l'équipe.

Des données de terrain moins formalisées sur les retours des enseignants-utilisateurs ont malgré tout pu être conservées par les chercheurs qui ont consigné les demandes et retours que les enseignants ont fait sur le fonctionnement de LNB. En outre, des débriefings ou d'autres échanges informels avec les enseignants et les responsables d'UE ont permis de recueillir des éléments pour comprendre leurs difficultés à concevoir des missions sur LNB et à utiliser LNB. L'ensemble de ces données a essentiellement servi les objectifs « technique » et « enseignement » du processus d'évaluation. Ces données non formalisées par un outil de mesure ne peuvent être exploitées dans un objectif de recherche.

Nous avons utilisé les traces d'activités des enseignants sur LNB en ciblant trois objectifs du PEL. Pour l'objectif « recherche », les traces permettent d'identifier les manières d'utiliser les fonctionnalités de LNB, ou de construire des profils d'utilisateurs de la plateforme. Pour les objectifs « enseignement » et « technique », les traces ont également été utilisées dans une visée de *learning analytics*. Ce travail a notamment permis la conception et l'implémentation de tableaux de bord sur l'interface enseignant de LNB. Pour l'objectif « institution », nous avons pu faire un retour aux financeurs sur l'utilisation effective de la plateforme par les enseignants et les étudiants concernés. L'utilisation de la plateforme dans l'enseignement a été croissant de 2017 à 2019. Dans cette période, le nombre d'enseignants a été multiplié par deux. Près de

4500 étudiants ont utilisé LNB. Plus de 8000 rapports d'étudiants ont été produits sur ces 3 années.

Après trois années d'évaluation, 1 345 questionnaires étudiants ont été recueillis, soit un taux de réponse de 26,8 % environ. Ce taux de réponse est conforme à ce qui peut être attendu dans la passation de ce type de questionnaire. Il conviendra cependant de vérifier que la répartition du taux de réponse soit uniforme entre les UE. Ces données fournissent des indicateurs pour la recherche (p. ex., perception d'un enseignement centré sur l'étudiant), pour l'enseignement (p. ex., utilisation et satisfaction) et pour l'institution (p. ex., satisfaction). La passation en ligne du questionnaire à la fin du cours s'est avérée peu efficace (peu de participation de la part des étudiants), ce qui nous a contraints à adopter une version papier du questionnaire. La saisie a été réalisée à l'aide d'un scanner des questionnaires remplis. Mettre en place ce type d'outil demande d'anticiper le recueil technique des données et de créer des alternatives en cas d'un faible taux de réponses

6.3. Bilan global

Le PEL et les outils fournis ont permis de répondre aux différents objectifs d'évaluation fixés au début du projet. Pour l>IDEX de l'université, le projet LNB a été le seul à fournir aux financeurs des données précises sur son utilisation et sur le degré de satisfaction des utilisateurs. En ce qui concerne l'enseignement et l'évolution technique, la formalisation d'entretiens, de débriefings et de séminaires a instauré un niveau de confiance et de collaboration entre les enseignants, l'équipe projet et les développeurs. Du côté recherche, les cadres théoriques mobilisés ont permis de travailler sur des indicateurs mesurés au fil du temps avec des outils adéquats (p. ex., guide d'entretiens, questionnaires). Le travail de formalisation des outils de mesures a produit des informations pour répondre aux questions de recherche initiales. En outre, le nombre important de données a permis d'élargir le champ des questions de recherche. Par exemple, des questions d'une granularité plus fine autour du travail collaboratif ont émergé.

De plus, le PEL a nécessité des aménagements au cours des passations. Les échanges par mail entre les chercheurs et les enseignants ont constitué des données précieuses pour faire évoluer LNB. De même, les débriefings réguliers avec les enseignants ont été importants pour, d'une part, contribuer aux améliorations de LNB et, d'autre part, maintenir la relation de collaboration autour de ce projet.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Aussi, le grand nombre de données produites et leur diversité peuvent être un écueil car il est parfois difficile d'analyser des corpus si divers. A ce sujet, une des difficultés rencontrées a été l'analyse des données demandant une compétence en analyse qualitative et quantitative. Bien que ces méthodes d'analyse soient maîtrisées par les deux méthodologues, les autres chercheurs du projet ont parfois eu du mal à adhérer aux résultats produits. Ils ont besoin d'analyser eux-mêmes les données pour s'approprier les résultats.

Enfin, travailler de manière longitudinale implique de garder un moyen de suivre les réponses d'une personne donnée. L'anonymisation est donc difficile. Il est nécessaire d'anticiper les procédures d'anonymisation des données et de mise en conformité avec la loi RGPD et les avis du comité d'éthique qui peuvent s'avérer longs. Dans notre cas, il a fallu un an de travail et d'aller-retour avant d'élaborer un document recevable par le responsable de la protection des données de l'Université.

7. Discussion et conclusion

Dans cette section, nous reviendrons d'abord sur les trois questions de recherche qui conduisent cet article. Puis, comme ce travail a été l'occasion de mettre en pratique les propriétés du *Design-Based Research*, nous réinterrogerons cette méthode quant à la conduite d'évaluation de plateformes en EIAH dans des contextes institutionnels.

Pour répondre à la 1^{ère} question « Quelles sont les particularités de l'évaluation des EIAH dans un contexte de recherche ? », nous avons proposé un cadre scientifique qui repose sur les travaux des sciences de l'artificiel, de manière à construire des connaissances scientifiques et des outils supports à cette connaissance. Nous avons aussi proposé une posture épistémologique, le constructivisme pragmatique, qui prend en compte les spécificités des EIAH évalués dans le cadre de projets de recherche financés et utilisés en contexte réel. Nous avons également choisi d'utiliser la méthode DBR en nous appuyant sur les sept propriétés posées par ses auteurs.

Pour répondre à la 2^e question « Quelles sont les méthodes de production de données adéquates ? », nous avons tout d'abord éclairci notre position sur des termes parfois mal mobilisés (qualitatif *versus* quantitatif, factuelle *versus* déclarative, etc.). Nous avons ensuite listé et décrit un

ensemble de méthodes de production des données envisageables en fonction des objectifs visés dans le cadre du projet LNB. Nous avons explicité en quoi le choix des méthodes de production ou de recueil des données est en lien avec les objectifs, les questions de recherche, la temporalité du projet, les indicateurs.

Pour répondre à la 3^e question « Comment construire un tel processus afin de produire des données pertinentes pour l'ensemble des acteurs du projet ? », nous avons proposé un modèle de processus découpé en six temps, eux-mêmes subdivisés en tâches. Ce modèle de processus a été appliqué au projet LNB conduit en contexte réel durant trois années.

Pour les plateformes développées dans le cadre des recherches en EIAH, nous avons proposé une structuration et une manière de conduire ces évaluations sur le long terme. La documentation de ces étapes et de leurs itérations (voir figure 2) peut être réutilisée et adaptée à d'autres contextes.

Notre travail a aussi été l'occasion de mettre à l'épreuve les sept propriétés du DBR. C'est sur ce point que nous souhaitons conclure, ouvrir le débat et proposer des réflexions sur les méthodes pour les recherches s'appuyant sur des plateformes numériques. Notre méthode a suivi les sept propriétés du DBR, nous souhaitons indiquer ici en quoi elles ont pu être suivies ou non.

Les propriétés « Pragmatique » et « Fondée sur le terrain » ont été appliquées sans difficulté, notre projet est effectivement réalisé en grandeur réelle avec des enseignants et des étudiants sur une durée de trois ans. Il en est de même pour la propriété « Intégrative » sur la mise en œuvre de plusieurs méthodes de production des données. Nous avons, en effet, utilisé au total sept méthodes de production de données, mesurant à la fois des données qualitatives et quantitatives et des données factuelles et déclaratives. Ce travail de construction d'outils de mesures et de méthodes de production de données est décrit dans une littérature abondante et de nombreux sites web répertorient des méthodes de production des données. Cependant, vu l'importance prise par les débriefings dans le projet LNB, une question est : comment tracer l'implicite, l'informel dans ce type de projet ?

Les propriétés « Itérative » de l'outil et « Flexible » de la question de recherche ont été tenues sans difficulté. La plateforme a été améliorée. De même, les questions de recherche ont évolué au fil du temps. La réflexion large sur les indicateurs au départ du projet et le nombre important de

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

données produites permettent aujourd'hui de répondre à d'autres questions de recherche de grain plus fin. Si les évolutions techniques de la plateforme sont tracées sur un dispositif numérique (p. ex., Github) ce n'est pas le cas de l'évolution des questions de recherche. Ainsi, nous pouvons nous interroger sur l'intérêt d'un outil destiné à suivre l'évolution des questions de recherche dans le cadre du DBR. Quelles fonctionnalités cet outil devrait-il offrir ? Comment devrait-il être structuré ?

Pour la propriété 4 « Interactive », nous étions dans un contexte favorable, car les enseignants chercheurs du projet sont eux-mêmes impliqués dans le développement informatique de LNB et dans la recherche didactique en sciences expérimentales. De plus, des méthodologues associés au projet ont apporté leur expertise dans la production des données. Sans ce contexte favorable, la propriété « interactive » du DBR aurait été difficile à mettre en place, car il s'agit de travailler de manière collaborative avec plusieurs acteurs ayant des intérêts et des contraintes parfois divergents. Pour le DBR, ces acteurs doivent travailler ensemble à l'élaboration de la recherche, des plateformes et des outils d'évaluation. Les domaines de compétences requis sont donc très différents. Dans ces conditions, faut-il mobiliser un outillage (Mandran, 2018) ? Ou bien avoir recours à un médiateur de type « *Broker* » (Sanchez et Monod-Ansaldi, 2015) pour conduire ces recherches ancrées dans le terrain ?

La propriété « Contextuelle » concerne la nécessité de faire le lien entre les résultats et le contexte dans lequel les données et les résultats ont été produits. Plusieurs contextes doivent être étudiés ; la multiplicité de contextes permet de garantir la généricité des résultats et d'augmenter leur niveau de validité. Dans notre cas, nous avons pu observer l'utilisation de la plateforme dans plusieurs contextes d'enseignement. Les résultats que nous avons produits sont pour l'instant au niveau global ou par unité d'enseignement en fonction des questions de recherche. Cette propriété pourra être validée par exemple par une étude comparative entre deux enseignements. Mais en quoi et comment pouvons-nous conclure que les résultats sont génériques ? Cette question ne peut être résolue qu'en se posant la question de la posture épistémologique adoptée pour construire une connaissance scientifique. C'est pourquoi notre ancrage dans le constructivisme pragmatique nous apparaît comme un

point essentiel avant même le choix d'une méthode de conduite de la recherche comme le *Design-Based Research*.

RÉFÉRENCES

Avenier, M.-J. (2019). Les Sciences de l'artificiel: une conceptualisation révolutionnaire de sciences fondamentales à parachever. *Projectics / Proyéctica / Projectique*, 24(3), 43-56. <https://doi.org/10.3917/proj.024.0043>

Avenier, M.-J. et Thomas, C. (2015). Finding one's way around various methodological guidelines for doing rigorous case studies: A comparison of four epistemological frameworks. *Systèmes d'information et Management*, 20(1), 61-98.

Balacheff, N. et Lund, K. (2013). Multidisciplinarity vs. Multivocality, the case of "learning analytics". Dans *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK'13)* (p. 5-13). <https://doi.org/10.1145/2460296.2460299>

Bernhaupt, R., Obrist, M., Weiss, A., Beck, E. et Tscheligi, M. (2008). Trends in the Living Room and Beyond: Results from Ethnographic Studies Using Creative and Playful Probing. *ACM Computers in Entertainment*, 6(1). <http://doi.acm.org/10.1145/1350843.1350848>

Brooke, J. (1996). SUS - A quick and dirty usability scale. Dans *Usability Evaluation in Industry* (p. 189-194). Taylor & Francis.

The Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8. <https://doi.org/10.3102/0013189X032001005>

Collins, A. (1992). Toward a Design Science of Education. Dans E. Scanlon et T. O'Shea (dir.), *New Directions in Educational Technology* (pp. 15-22). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77750-9_2

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

De Vries, E. et Baillé, J. (2006). Apprentissage: Référents théoriques pour les EIAH. Dans M. Grandbastien et J.-M. Labat (dir.), *Environnements Informatiques pour l'apprentissage humain* (p. 27-45). Hermes Science.

Donnelly, D. F., Linn, M. C. et Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research*, 84(4), 572-608.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. et Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Dans *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.

Granić, A. et Ćukušić, M. (2011). Usability testing and expert inspections complemented by educational evaluation: A case study of an e-learning platform. *Journal of Educational Technology & Society*, 14(2), 107-123.

Hannafin, M. J. et Hannafin, K. M. (2010). Cognition and Student-Centered, Web-Based Learning: Issues and Implications for Research and Theory. Dans J. M. Spector, D. Ifenthaler, P. Isaias, Kinshuk et D. Sampson (dir.), *Learning and Instruction in the Digital Age* (p. 11-23). Springer US.

**Nadine MANDRAN, Maelle PLANCHE, Patricia MARZIN-JANVIER,
Mathieu VERMEULEN, Aous KAROUI, Cédric d'HAM, Isabelle
GIRAULT, Claire WAJEMAN, Christian HOFFMANN**

Hannafin, M., Land, S. et Oliver, K. (1999). Open Learning Environments: Foundations, methods, and models. Dans C. Reigeluth (dir.), *Instructional-Design Theories and Models* (Vol. 2, p. 115-140). Lawrence Erlbaum.

Herrington, J., McKenney, S., Reeves, T. et Oliver, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. *EdMedia+ Innovate Learning*, 4089-4097.

Hoffmann, C., Girault, I., Kahane, C., d'Ham, C. et Planche, M. (2019). Utilisation d'une plateforme numérique dans un dispositif d'apprentissage par problèmes (APP). QPES, Brest, France.

Hoffmann, C., Girault, I., Kahane, C., d'Ham, C. et Planche, M. (2021). Introduction d'une plateforme numérique dans un dispositif APP - quelles adaptations, quelle influence sur la coopération/collaboration étudiante? Les Annales de QPES, 1(3), Article 3. <https://doi.org/10.14428/qpes.v1i3.62083>

Land, S. et Oliver, K. (2012). *Open Learning Environments*. Dans N. M. Seel (dir.), *Encyclopedia of the Sciences of Learning* (p. 855). <https://doi.org/10.1007/978-1-4419-1428-6>

Lathika, K. (2016). Student centered learning. *International Journal of Current Research and modern education*, 677-680.

Lee, E. et Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: Own it, learn it, and share it. *Educational Technology Research and Development*, 64(4), 707-734.

Lhoste, Y. et Orange, C. (2015). Quels cadres théoriques et méthodologiques pour quelles recherches en didactique des sciences et des technologies? *R DST. Recherches en didactique des sciences et des technologies*, 11, 9-24.

Mandran, N. (2018). *Traceable human experiment design research: Theoretical model and practical guide* (ISTE, Vol. 9). Wiley-ISTE.

Mandran, N. et Dupuy-Chessa, S. (2017). THEDRE: A Traceable Process for High Quality in Human Centred Computer Science Research. Dans *Proceedings of the International Conference on Information Systems Development (ISD 2017)*. <https://aisel.aisnet.org/isd2014/proceedings2017/ISDMethodologies/12>

Mandran, N., Vermeulen, M. et Prior, E. (2021). Comment guider les doctorants dans l'utilisation du Design-Based Research? Dans *Actes de la 10^e Conférence Environnements Informatiques pour l'Apprentissage Humain (EIAH 2021)* (p. 130-141). <https://hal.archives-ouvertes.fr/hal-03292805>

McKenney, S. et Reeves, T. C. (2019). *Conducting educational design research* (second). Routledge.

Nielsen, J. (1994). *Usability Engineering*. Elsevier.

Paillet, P. et Mucchielli, A. (2011). *L'analyse qualitative en sciences humaine et sociales* (2^e éd.). Armand Colin.

Parahoo, S. K., Santally, M. I., Rajabalee, Y. et Harvey, H. L. (2016). Designing a predictive model of student satisfaction in online learning. *Journal of Marketing for Higher Education*, 26(1), 1-19. <https://doi.org/10.1080/08841241.2015.1083511>

Planche, M., d'Ham, C., Hoffmann, C., Mandran, N., Girault, I., Wajeman, C. et Marzin, P. (2021). Soutien aux pédagogies actives et transformations impulsées par

une plateforme numérique dédiée aux sciences expérimentales. Dans *Actes de la 10^e Conférence Environnements Informatiques pour l'Apprentissage Humain (EIAH 2021)*. <https://hal.archives-ouvertes.fr/hal-03293473>

Planche, M., Mandran, N., Girault, I., Marzin-Janvier, P., d'Ham, C. et Wajeman, C. (2019). Contribution de différents outils de mesure à l'évaluation des usages d'une plateforme numérique par un processus longitudinal: Cas du travail à distance. Dans *Actes de la 9^e Conférence "Environnements Informatiques pour l'Apprentissage Humain (EIAH 2019)*.

Ray, D. (2001). *Mesurer et développer la satisfaction des clients*. Editions d'Organisation.

Rentroia-Bonito, M. A. et Jorge, J. A. P. (2003). An Integrated Courseware Usability Evaluation Method. Dans V. Palade, R. J. Howlett et L. Jain (dir.), *Knowledge-Based Intelligent Information and Engineering Systems* (Vol. 2774, p. 208-214). Springer. https://doi.org/10.1007/978-3-540-45226-3_29

Sanchez, É. et Monod-Ansaldi, R. (2015). Recherche collaborative orientée par la conception. *Éducation et Didactique*, 9(2), 73-94.

Simon, H. A. (2004). *Les Sciences de l'artificiel* [The sciences of the artificial] (J.-L. Le Moigne, trad.). Gallimard. (Ouvrage original publié en 1996, 3^e éd. rev. et complétée).

Tiberghien, A., Vince, J. et Gaidioz, P. (2009). Design-based Research: Case of a teaching sequence on mechanics. *International Journal of Science Education*, 31(17), 2275-2314. <https://doi.org/10.1080/09500690902874894>

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner et E. Soubberman, dir.). Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>

Wang, F. et Hannafin, M. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4), 5-23.