# UNE RESSOURCE DE FORMATION BASÉE SUR UNE EXPLORATION DU HASARD AU CYCLE 3

**Jannick TRUNKENWALD**
Professeur Agrégé, Lycée International Alexandre Dumas (LIAD, Alger)
Laboratoire de Didactique André Revuz (LDAR, Université de Paris)
jannick.trunkenwald@ac-paris.fr

**Abderahim BOUTAHAR**
Référent Mathématique de Circonscription, IEN Lens (Académie de Lille)
abderahim.boutahar@ac-lille.fr

**Djalila BENYAHIA**
Professeure des Écoles, École Internationale Alexandre Dumas (EPIAD, Alger)
Laboratoire de Mathématiques Maurice Audin (LIAD, Alger)
djalila.benyahia@liad-alger.fr

**Anissa IDJER**
Professeure de Mathématiques, Lycée International Alexandre Dumas (LIAD, Alger)
Laboratoire de Mathématiques Maurice Audin (LIAD, Alger)
anissa.idjer@liad-alger.fr

**Slimane KADDOUR**
Professeur des Écoles, Petite École Hydra (PEH, Alger)
Laboratoire de Mathématiques Maurice Audin (LIAD, Alger)
slimane.kaddour@mlfmonde.org

**Résumé**
L'implémentation au Lycée International Alexandre Dumas (Alger) d'un laboratoire de mathématiques (mesure 16 du plan Villani-Torossian) a permis l'intégration d'expérimentations filmées en classe, au dispositif de formations d'un réseau d'établissements à Alger. L'objectif de réaliser une ressource portant sur la modélisation mathématique, a entraîné une série d'analyses de séances désormais exploitables dans la cadre du plan de formation des lycées français de la zone Maghreb-Est. La séquence a été expérimentée en CM1, CM2, et 6ème. Elle porte sur l'usage d'outils de gestion de données dans une situation confrontant les élèves aux fluctuations du hasard. L'analyse des séances exploite le concept d'Espace de Travail Mathématique afin de mieux identifier le processus de validation. Il en ressort différentes observations spécifiques au terrain du cycle 3, qui correspondent aux dimensions instrumentale, sémiotique, et discursive du travail mathématique réalisé par les élèves.

## I - UNE INGÉNIERIE POUR LA FORMATION DE FORMATEUR

### 1 Contexte institutionnel

Le Lycée International Alexandre Dumas (LIAD) est un établissement scolaire français établi sur 4 sites à Alger (Ben Aknoun et Dely Brahim), Oran, et Anaba, qui est directement géré par l'Agence pour l'Enseignement Français à l'Étranger (AEFE) depuis Paris. Le LIAD est aussi l'unique centre d'examen pour tous les élèves algériens, ou résidents en Algérie, qui passent les épreuves du brevet des collèges et du baccalauréat français.

#### 1.1 *Un dispositif de développement des pratiques professionnelles*

Le LIAD bénéficie d'un dispositif de formation qui est associé à la zone Maghreb-Est des lycées français de l'AEFE (MEM). Un Plan de Formation Continue est élaboré chaque année à partir des remontées de besoins exprimés par les coordonnateurs disciplinaires de tous les lycées homologués de la zone de formation. Des cellules de formation continue transmettent alors une liste restreinte de propositions de formation aux personnels chargés d'une mission de conseil pédagogique, qui élaborent un cahier des charges applicable sous la forme de stages animés l'année suivante.

Par ailleurs en décembre 2019, le Laboratoire de Mathématiques Audin a été inauguré au LIAD par l'Ambassade de France. Il s'agit d'une structure d'accueil et de développement des pratiques professionnelles au sens de la mesure 16 de Plan Villani-Torossian, qui organise divers travaux liés à l'enseignement des mathématiques au sein d'axes de réflexions. Ce laboratoire intervient aussi dans le domaine de la coopération éducative en associant à ses travaux une douzaine d'établissements scolaires algériens. Des séminaires périodiques permettent de mettre en valeur certains chantiers pédagogiques liés aux mathématiques, et aussi de bénéficier d'intervenants issus d'universités partenaires. De plus au niveau de l' « axe didactique » du laboratoire, un collectif d'enseignants est engagé dans le parcours Master 2 recherche en Didactique des Mathématiques encadré par l'Université d'Aix-Marseille.

Ce contexte permet d'envisager une interaction du Laboratoire Audin avec le dispositif de formation de la zone MEM. Cela peut prendre la forme d'une activité de veille pédagogique, en appui sur les outils issus de la recherche en didactique des mathématiques. Il peut aussi s'agir de concevoir des ressources pédagogiques destinées à la formation, ou encore de tester, voir reconditionner, certains outils ou activités proposés lors de stages. Mais la formation n'étant pas d'un point de vue institutionnel dans les prérogatives directes du laboratoire, l'action envisagée doit plutôt s'inscrire dans le déroulement d'un projet pédagogique.

Par ailleurs, il est désormais pertinent, après ces deux premières années de fonctionnement du Laboratoire Audin, de rechercher une forme de validation au niveau de l'institution du LIAD, du principe associé à ce type de « recherche action ». Cela permet aussi de s'inscrire, après 2 années de fonctionnement de la structure d'accueil, dans une démarche d'évaluation au niveau du Service de Coopération et d'Action Culturelle de l'Ambassade de France, et du réseau des laboratoires de mathématiques de l'AEFE. Il s'agit de l'un des éléments de motivations ayant précédé le lancement du projet pédagogique présenté dans cet article. Un dernier objectif correspond à la constitution par le laboratoire Audin d'un vivier de formateurs potentiels pour l'enseignement des mathématiques, qui peuvent à la fois intervenir dans le domaine de la coopération éducative auprès des établissements partenaires, et aussi intervenir dans le dispositif MAT-PAT (Maîtres et Professeurs d'Accueil Temporaire) qui ouvrent leurs classes pour appuyer le dispositif de formation initiale destiné aux nouveaux recrutés enseignants PDL (Personnels de Droit Local).

### 1.2 *Un cahier des charges lié à l'institution scolaire*

La conception par le Laboratoire de Mathématiques Audin d'un outil de formation doit à la fois répondre à certaines attentes institutionnelles, et viser un changement de paradigme au niveau de certaines pratiques professionnelles. Pour cela, il est pertinent pour la structure d'accueil de s'appuyer sur sa propre expertise ciblée dans le domaine de la didactique des mathématiques.

Le présent projet s'inscrit dans le cadre du conseil école-collège du LIAD, en liaison avec les établissements primaires français à Alger qui sont l'École Internationale Alexandre Dumas (EPIAD) et la Petite École Hydra (PEH). L'objectif pédagogique est lié à l'approche scientifique au niveau du cycle 3, en lien avec le développement de compétences transversales telles que la modélisation, la représentation des données, et la démarche de recherche. Le conseil école-collège attend de plus que ce projet pédagogique permette une interaction entre des élèves de différents niveaux de classes.

### 1.3 *Des objectifs de recherche associés au projet pédagogique*

Nous avons donc décidé de baser la conception de cette formation sur des travaux expérimentaux menés au LIAD, dans le cadre des activités du laboratoire Audin, au niveau de classes du lycée. Ces travaux participent de la phase pré-expérimentale d'une thèse de didactique des probabilités en cours : « Approche fréquentiste et approche combinatoire de la notion de probabilité. Nature du travail mathématique mis en jeu. » En particulier, une séance expérimentale menée en classe de seconde doit, dans le cadre d'un élément du programme officiel, mener les élèves à découvrir la notion de probabilité. Il s'agit d'une approche de type fréquentiste basée sur la production d'échantillons de même taille dont les fréquences de succès associées fluctuent autour de la valeur de probabilité. Les observations menées à partir de cette expérience ont permis d'identifier certains aspects du travail réalisé en classe par les élèves, suivant trois phases de traitement de type empirique, qui sont successivement : tester l'expérience aléatoire, construire un échantillon, produire une liste de fréquences simulées à partir de plusieurs échantillons de même taille.

Les résultats observés ne permettaient cependant pas, à ce niveau de classe, d'appréhender suffisamment certains enjeux d'ordre cognitifs liés à l'émergence de cette perception rationaliste du hasard qui correspond à la notion de probabilité. Nous pensons en effet que soumettre à des élèves de niveau primaire une telle situation probabiliste, permet d'observer davantage les interactions entre les différents domaines mathématiques en jeu dans cette approche fréquentiste tels que : le nombre, la gestion de données, la géométrie, le calcul numérique.

Par ailleurs, nous faisons l'hypothèse que la notion de probabilité ne fait pas encore partie, au niveau du primaire, de l'environnement culturel habituel des élèves, ce qui est justement intéressant du point de vue d'une étude portant sur l'introduction de la notion de probabilité.

Le programme officiel n'aborde pas la notion de probabilité à ce niveau du cycle 3. Ce contexte institutionnel nous amène à fixer pour l'élaboration de cette formation un objectif central lié au domaine mathématique de la gestion et organisation de données. Une situation liée au hasard, mettant en scène une expérience aléatoire, doit alors être abordée, à travers une approche empirique, de type scientifique. Cela doit alors amener les élèves à exploiter ce domaine de la gestion de donnée, à l'aide aussi des connaissances de base portant sur le nombre, le calcul, et les outils géométriques de représentation.

Nous insistons toutefois sur cette idée que la question liée à l'introduction de la notion de probabilité analysée ici à l'éclairage d'éléments issus de la recherche en didactique des mathématiques, n'est pas l'objet de la formation mise en place. Nous souhaitons simplement initier une synergie entre le dispositif de formation et certaines activités de recherche menées au sein du Laboratoire Audin.

L'objectif du projet pédagogique présenté dans cet article, est finalement de concevoir une formation de formateurs permettant de sensibiliser à une approche constructiviste de l'enseignement des mathématiques, qui s'appuie sur la démarche d'investigation. Nous présenterons aussi un retour d'impression de ce qu'a pu apporter une telle ingénierie collaborative entre chercheur et enseignants, dans une perspective de formation. Le travail de recherche sous-jacent étant focalisé sur l'élève en tant que sujet cognitif, nous avons en effet associé les enseignants expérimentateurs à cet enjeu en les sensibilisant au cadre théorique et à la méthodologie.

## 2 Cadre théorique de l'étude didactique

Nous présentons dans ce paragraphe quelques outils issus de la recherche en didactique des mathématiques, auxquels les enseignants expérimentateurs engagés dans ce projet pédagogique ont été sensibilisés. L'objectif était de pouvoir mettre des mots sur certains observables, et sur certaines pratiques en classe, et aussi d'expliciter l'intérêt des expérimentations du point de vue de la conception d'un outil de formation destiné à la zone MEM, en amenant ces enseignants à réfléchir à leurs propres pratiques au cours de l'ensemble de cette expérimentation.

### 2.1 *Les cycles de modélisations*

La modélisation s'invite très naturellement dans le lien entre le monde réel, abordé à travers le travail d'observation statistique de données brutes ou de phénomènes aléatoires, et le monde probabiliste, constitué d'évènements abstraits dont la mesure est bien fixée. C'est bien l'idée de modèle mathématique qui rend possible cette dialectique entre les pensées empirique et rationaliste. Rappelons ce que nous entendons par modélisation :

> *Parmi les différents registres de représentations, le langage et le symbolisme mathématique permettent des descriptions puissantes sur lesquelles peuvent opérer des propriétés et des algorithmes généraux. Nous les appellerons « modèles mathématiques » (Henry, 1999, p. 26).*

Et dans le contexte qui nous intéresse, le rôle de la modélisation est renforcé : l'enseignement des probabilités commence à peine au cours des études secondaires ce qui en fait une discipline d'éveil dont les fondements théoriques sont balbutiants. Cela nécessite de rester proche du monde concret et entraîne la démarche de modélisation. Dans les domaines déjà abordés dès le début du primaire, tels que le calcul ou la géométrie, le statut de modèle mathématique des différentes notions s'est estompé avec la pratique et la mise en place d'expertises fournissant un champ d'activité en interne du modèle (au sein des mathématiques). Du côté des probabilités telles qu'abordées par l'institution scolaire au début du lycée, il reste difficile intuitivement d'accepter une mesure du hasard et le référentiel théorique ne fournit pas non plus, à ce stade de la scolarité, un champ d'application conséquent qui soit interne au modèle probabiliste. Nous pouvons d'ailleurs remarquer que les probabilités sont un des seuls domaines des mathématiques où on reste autant intéressé à la réalité…. On le remarque aisément en observant la forme de la plupart des énoncés de baccalauréat portant sur le thème des probabilités. Et c'est en cela que l'aspect de modélisation prend tout son sens.

Dans le cadre de cette étude, il est donc essentiel de pouvoir expliciter les enjeux épistémologiques et cognitifs associés à la démarche empirique qui permet d'étudier un phénomène aléatoire. Cette approche fréquentiste peut être pratiquée en amont de la découverte de la notion de probabilité et ceci dès que des notions de gestion des données sont abordées par les élèves. Il peut s'agir d'une démarche en trois phases consistant à tester l'expérience aléatoire, puis répéter cela afin de constituer un échantillon permettant de quantifier un nombre de succès obtenus, et enfin de répéter plusieurs fois ce protocole pour produire plusieurs échantillons de même taille qui permettront d'observer le phénomène de fluctuation des fréquences obtenues.

Chacune de ces trois phases correspond à un travail de modélisation dans le domaine des statistiques descriptives, suivi d'une interprétation dans le domaine réel du résultat mathématique obtenu. Nechache (2016, p. 204) présente dans sa thèse une version synthétique de description de tels cycles de modélisation qui est découpée en 3 principales étapes :

Étape 1 « Description de la réalité » : Il s'agit de décrire la situation réelle dans le langage courant et de la traduire « en système simplifié et structuré : c'est le niveau du modèle probabiliste, donné en termes pseudo-concrets »(Henry,1999,p. 29) ;

Étape 2 « Mathématisation » : Il est question de « la mathématisation ou le formalisme du modèle »(Ibid.,p.29). Autrement dit, on représente le modèle réel dans la symbolique propre aux mathématiques (mise en équations et leurs résolutions, en utilisant des outils adaptés, par exemple) ;

Étape 3 « Validation » : Cette étape porte sur la validation du modèle choisi. Il s'agit d'interpréter les résultats obtenus dans le modèle mathématique afin de fournir une réponse au problème posé dans la situation réelle.

Nous devons donc envisager un tel découpage de l'activité des élèves suivant trois cycles de modélisation successifs comportant chacun les trois étapes précitées.

### 2.2 *L'Espace de Travail Mathématique*

Nous partons du constat fait par Nechache (2016) qui remarque l'absence au début du lycée *d'un référentiel théorique probabiliste qui permettrait d'aborder cet intervalle de fluctuation par un raisonnement de type déductif.* L'approche de cet intervalle de fluctuation reste en effet, à ce stade, expérimentale et basée sur l'observation des fréquences d'échantillons de même taille. Chaque valeur de fréquence correspond alors à un échantillon obtenu lui-même en répétant un certain nombre de fois l'expérience aléatoire considérée. Il s'agit bien d'une démarche empirique basée sur un raisonnement de type inductif. Afin de cerner au-mieux les éléments du référentiel théorique opérant au cours du travail mathématique mené par les élèves dans chacun de ces deux domaines, nous allons devoir identifier leur rôle dans le processus de validation de ce travail par une preuve. Pour cela nous devrons prendre en considération les signes et symboles permettant un traitement mathématique au sein de chaque système de représentation, ainsi que la conversion du travail mathématique d'un registre sémiotiques mis en œuvre, à l'autre. Enfin, en lien avec l'ensemble de ce processus nous devrons identifier et distinguer la dimension instrumentale basée sur l'usage de schèmes d'actions, de schèmes mentaux, ou de boites noires, que l'on nommera artefacts. Cette dimension instrumentale n'appartient en effet ni à la discipline des mathématiques en elle-même, ni à son expression visuelle, mais participe du travail mathématique en contribuant à la connaissance. Nous sommes de ce point de vue amenés à exploiter l'Espace de Travail Mathématique (ETM) tel qu'il a été défini par Kuzniak (2011). Les outils technologiques (*artefacts*), sémiotiques (*representamen*), et théoriques (*référentiel*) du plan épistémologique, peuvent alors être exploités à l'aide de schèmes appropriés. On parle dans ce cas respectivement de *genèses instrumentale*, *sémiotique*, et *discursive*. Ces genèses respectives produisent dans le plan cognitif des *constructions*, des *visualisations*, et des *preuves*. Enfin, l'activation de deux genèses peut entraîner une circulation entre les dimensions qui les portent, ce qui peut parfois être rapproché de l'idée de modélisation.

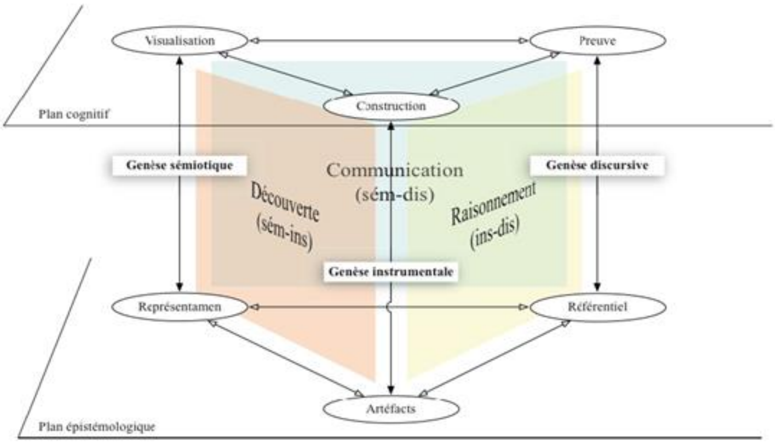



*Figure 1. L'Espace de Travail Mathématique (Kuzniak & Richard, 2014)*

## 3 Méthodologie d'étude du travail des élèves

Pour rappel, la notion de probabilité n'est pas abordée avec les élèves de cycle 3 dans le cadre de ce projet pédagogique. Les enseignants expérimentateurs sont cependant sensibilisés à l'intérêt de mener en classe une exploration du hasard à l'aide des outils de gestion de données. Cet intérêt est en effet double, du point de vue de l'activité mathématique menée en classe dans le domaine statistique, et aussi de point de vue des prolongements ultérieurs de ce travail qui seront menés par les élèves au niveau du cycle 4 pour aborder la notion de probabilité.

### 3.1 *La problématique en jeu*

Pour établir l'intervalle de fluctuation en 2$^{nde}$, l'absence de justification formelle disponible dans le référentiel théorique des élèves a été abordée par Parzysz (2009) :

> *L'accès à la notion de modèle qui est une finalité visée à terme par le cycle terminal, peut être préparé par l'étude et la simulation d'expériences aléatoires diverses correspondants au même modèle probabiliste. La comparaison des procédures, des tableaux et des hypothèses sous-jacentes doit permettre aux élèves de se convaincre de l'analogie que présentent ces expériences sous leurs apparences diverses, et de déboucher sur la notion de schéma d'expérience, constituant en quelque sorte une classe d'équivalence d'expériences aléatoires. (Parzysz, 2009, p. 102).*

Nechache (2016) aborde aussi ces questions en présentant le cycle de modélisation d'un modèle mathématique de type numérique, l'expérience réelle est d'abord simplifiée sous forme d'une expérience aléatoire permettant d'envisager un premier modèle probabiliste réel. Celui-ci est alors éventuellement traduit en modèle numérique pour la simulation. L'exécution de ce programme de simulation peut donner une réponse, qui est d'abord interprété par rapport au modèle réel, puis interprété en regard de l'expérience aléatoire. Nechache souligne cependant une difficulté liée à ce type de modèle numérique :

> *Dans l'enseignement secondaire, la construction du modèle probabiliste de type numérique est basée sur des connaissances qui ne sont pas disponibles dans l'espace de travail personnel des élèves. C'est pourquoi la construction de ce modèle est habituellement prise en charge par le professeur qui laisse aux élèves l'exécution de la simulation. (Nechache, 2016).*

Nous souhaitons compléter cette réflexion liée à la problématique spécifique d'une réalisation du programme de simulation qui serait conjointe à son exploitation d'un point de vue probabiliste. Nous formulons l'hypothèse que : « le modèle final simulant la fluctuation d'échantillonnage est élaboré dans de meilleures conditions si on élabore en classe plusieurs cycles de modélisations abordant successivement l'expérience aléatoire, la construction d'un échantillon, puis la représentation des fréquences de succès associées à une liste d'échantillons de même taille. »
Nous adaptons cependant ce questionnement au contexte institutionnel d'un niveau de classe où la notion de probabilité n'est pas au programme. Même s'il est envisagé d'aller ultérieurement plus loin avec une exploitation du logiciel Scratch, nous restons dans un premier temps, et dans le cadre de cette communication, focalisés sur une expérience réelle, non simulée par l'informatique. Cela doit nous permettre d'observer le travail mathématique réalisé par les élèves dans le cadre d'une première rencontre menant à une exploitation des mathématiques pour aborder le hasard.
En lien avec la phase pré-expérimentale associée au sujet de thèse en cours portant sur l'introduction de la notion de probabilité, nous identifions donc une question visant à apporter un complément d'éclairage sur les enjeux épistémologiques et cognitifs de l'approche fréquentiste d'une fluctuation d'échantillonnage, au niveau du cycle 3 : « Quelle est la nature du travail mathématique mis en œuvre par des élèves de cycle 3, lorsqu'ils mobilisent les outils de gestion de données pour étudier une expérience aléatoire en produisant des échantillons de même taille ? »
Le choix de la tâche se base sur l'idée d'une situation mathématique classique qui présente toutefois les caractéristiques d'une tâche riche pour les élèves. Afin de préserver un intérêt pour le travail empirique associé à l'approche fréquentiste, nous évitons une situation probabiliste qui induirait trop facilement un raisonnement basé sur l'équiprobabilité, mais nous en préservons la possibilité.

Énoncé : JEU A : On lance un dé à 6 faces. On gagne si on obtient au-moins 5.

JEU B : On lance deux dés à 6 faces. On gagne si la somme des résultats vaut au-moins 9. Pour quel jeu est-il le plus facile de gagner ?

### 3.2 *Le déroulement prévu pour l'expérimentation initiale (CM2)*

L'expérimentation réalisée en 2019/20 s'est déroulée dans une classe de CM2 de l'École Internationale Alexandre Dumas (EPIAD) établissement du réseau AEFE à Alger. Elle a été interrompue par la crise du COVID juste avant la séance 5 qui devait permettre de représenter sur un graphique les nombres de succès correspondants aux gros échantillons (obtenus après regroupement des échantillons construits par les binômes, voir détails dans la suite de l'article).

•Séance 1 (21/11/2019) : **Découverte de deux jeux de hasard.**

Présentation et explicitation de l'énoncé du problème Les élèves manipulent les dés et relèvent leurs données librement. Mise en commun des observations menées. Conclusion sur la question posée.

•Séance 2 (09/12/2019) : **Comparaison à partir de petits échantillons standards**

Une discussion avec la classe doit lancer l'idée de construire des tableaux pour mieux organiser les données de 20 lancers pour chaque jeu. Les élèves testent les deux jeux et rassemblent leurs résultats dans des tableaux qu'ils construisent librement. La discussion doit ensuite mener à une insuffisance d'arguments pour conclure sur la question posée. L'objectif est de parvenir en fin de séance à l'idée de tableaux standards afin de pouvoir comparer ou regrouper le contenu de plusieurs tableaux.

•Séance 3 (23/01/2020) : **Réalisation par binôme de 5 échantillons de taille 20 pour le jeu B.**

La réalisation de ces 5 tableaux relevant chacun les sommes obtenues pour 20 lancers des deux dés, et dénombrant les succès, est un travail attendu de la part des élèves. Ils ont été préparés à cela lors de séance précédente. Les élèves étaient attendus sur la manière dont ils allaient notifier les bilans de chaque tableau.

Une discussion doit mener à un moyen d'obtenir beaucoup plus de données. Finalement, tous les tableaux sont mis en commun sous forme d'un grand tableau commun présentant 5 lignes et 11 colonnes. Les élèves sont attendus sur la manière d'exploiter en termes statistiques ce grand tableau.

•Séance 4 (30/01/2020) :**Réalisation d'une représentation graphique des données obtenues.**

Dans un premier temps une représentation libre de leurs données de binôme. Des discussions doivent porter sur la manière de représenter. Dans un deuxième temps une représentation standardisée (axes et échelle) est attendue.

### 3.3 *Une reprise de l'expérimentation en 2020/21 (6ème , CM2 , CM1)*

L'ensemble de ce parcours expérimental a été reconduit en 2020/21 au niveau d'une classe de CM2 de la Petite École Hydra (PEH) établissement du réseau de la Mission Laïque Française (MLF) à Alger. Cette deuxième vague d'expérimentations en CM2 a permis d'aller plus loin en réalisant aussi une séance intermédiaire dédiée à la réflexion menant au regroupement des données d'échantillons correspondant à tous les binômes : un document préparé par l'enseignant présentait le tableau de synthèse avec en plus un graphique de synthèse montrant tous les échantillons de tous les groupes. Cela devait faciliter et enrichir la réflexion menant aux gros échantillons.

Deux autres expérimentations analogues ont permis de mener cette expérimentation dans une classe de 6ème du LIAD et dans une classe de CM1 de l'EPIAD. Un bilan d'étape a pu se dérouler en visioconférence avec ces trois classes, afin de confronter les idées des élèves sous forme d'un petit séminaire de recherche.

La phase envisagée d'exploitation de l'informatique à l'appui du logiciel Scratch n'a cependant pas pu être mise en œuvre en raison d'un manque de temps et de calendrier.

## II - ANALYSE A PRIORI DU TRAVAIL DES ÉLÈVES

Du point de vue mathématique, la séquence peut être découpée ainsi : modélisation statistique de l'expérience aléatoire, modélisation statistique de l'échantillon, et modélisation statistique de la fluctuation des fréquences.

Chacune de ces trois phases de l'expérimentation correspond à un cycle de modélisation. Pour chacun de ces cycles, nous décrivons de manière succincte les différents aspects du travail mathématique avec : 1) les étapes de validation mathématique avec les éléments du référentiel théorique, 2) les étapes de construction instrumentale avec les artefacts employés, 3) les étapes liées à la visualisation, avec les signes, symboles, et registres sémiotiques (et leurs phases de traitement et de conversion).

Nous accompagnons cette description d'extraits correspondants à une analyse plus générale du travail mathématique lié à ces trois cycles de modélisation, lorsqu'une simulation d'échantillonnage est mise en œuvre pour aborder de manière fréquentiste la notion de probabilité.

## 1 Tester l'expérience aléatoire (séance 1)

Ce premier cycle de modélisation se traduit en trois étapes :

La situation, dite réelle et correspondant à la comparaison de deux règles de jeux dans l'énoncé, est tout d'abord simplifiée sous la forme du scénario précis correspondant à l'un de ces deux jeux. La réflexion peut alors par exemple se focaliser sur le jeu B.

Un schéma d'expérience est identifié au niveau pseudo-concret du processus de modélisation. Il s'agit aussi pour le jeu B d'appréhender les rôles respectifs des deux dés en regard des différentes issues possibles et de leur propension à se réaliser, et d'identifier les différents résultats menant à un succès.

L'étape de mathématisation doit ensuite être mise en œuvre. Relativement triviale, elle se situe dans le domaine des statistiques descriptives. Cela mobilise les notions d'expérience aléatoire et d'évènement. Il s'agit en particulier pour l'élève de comprendre que l'expérience est reproductible à l'identique, que les résultats obtenus s'expriment dans un univers ensembliste constituant des évènements, et qu'il est envisagé de mesurer des chances de réussite de cette même expérience.

Dans le cas où l'expérience est simulée à l'aide d'un outil informatique, le travail correspondant se complexifie mais dans le cadre de la présente expérimentation l'usage de l'informatique n'a pas été mis en œuvre. Cette étape de mathématisation correspond alors à un registre de représentation particulier qui donne aussi lieu à un travail de nature syntaxique ou nécessitant l'emploi de schèmes spécifiques à l'outil employé.

Les réponses apportées par le modèle mathématique sont enfin interprétées en regard du modèle pseudo-concret lié au schéma d'expérience. Les élèves sont amenés à constater la cohérence éventuelle entre le modèle probabiliste envisagé et les résultats du schéma d'expérience testé. C'est à ce stade que les élèves peuvent envisager un nouveau cycle de modélisation mobilisant le domaine probabiliste pour interpréter le phénomène par une approche de type combinatoire. L'idée de tester un certain nombre de fois l'expérience aléatoire peut aussi émerger à cette étape du processus de modélisation, ce qui relance un cycle de modélisation en mobilisant le domaine des statistiques descriptives.

*a) Premier cycle de modélisation : Simuler l'expérience aléatoire*

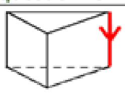
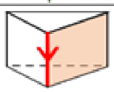
<table>
<tr><th>Phases</th><th colspan="2">ESPACE DE TRAVAIL MATHEMATIQUE</th><th>Outil mobilisé</th><th>Objet en construction</th></tr>
<tr><td rowspan="3">Etape 1</td><td>Plan cognitif</td><td>Lecture de l'énoncé et entrée dans le questionnement.</td><td rowspan="3">A partir du référentiel : La notion théorique d'expérience aléatoire.</td><td rowspan="3">Identification de l'expérience aléatoire correspondant à la tâche, avec l'évènement visé.</td></tr>
<tr><td>Circulation du travail mathématique et domaine en jeu</td><td> DIS ↓ (référentiel) Domaine : PROBA</td></tr>
<tr><td>Plan épistémologique</td><td>Mobilisation de la notion d'expérience aléatoire.</td></tr>
<tr><td rowspan="3">Etape 2</td><td>Plan cognitif</td><td>Manipulation des dés. Schèmes gestuels.</td><td rowspan="3">Utilisation de dés physiques à 6 faces (ou bien d'une instruction informatique du type Alea).</td><td rowspan="3">Construction d'une simulation réelle (ou informatique) de l'expérience aléatoire. Les résultats envoyés par les 3 dés lancés interprété comme succès ou échec.</td></tr>
<tr><td>Circulation du travail mathématique et domaine en jeu</td><td> DIS – INS ↓ (Artefact) Domaine : PROBA</td></tr>
<tr><td>Plan épistémologique</td><td>Mobilisation des artefacts (dés à 6 faces).</td></tr>
</table>

*Figure 2. Description de l'Espace de Travail Mathématique correspondant au premier cycle de modélisation*

Au niveau de l'étape de mathématisation, nous pouvons analyser l'ETM statistique mis en œuvre. Du point de vue de la dimension discursive, on pense à la mobilisation à partir du référentiel théorique des notions d'expérience aléatoire et de succès. Du point de vue de la dimension instrumentale, l'artefact « dé » est mobilisé pour construire un test, ou une simulation, de l'expérience aléatoire.

La dimension sémiotique qui semble a priori peu sollicitée à ce stade de l'activité, mérite cependant d'être questionnée dans le cadre de notre recherche. Enfin, à l'issue de ce cycle de modélisation, le schéma d'expérience associé à l'expérience aléatoire considérée peut intégrer le plan épistémologique de l'ETM statistique en tant que nouvel artefact.

## 2 Construire empiriquement et interpréter un échantillon (séances 2 et 3)

Le jeu A présentant un caractère trivial, du point de vue de l'évaluation des chances de réussites, il a été prévu dans le méthodologie de l'étude d'orienter après une première phase de l'activité des élèves permettant la confrontation des deux jeux, d'orienter la réflexion des élèves sur le jeu B. La stratégie visant à répéter l'expérience aléatoire de manière à construire un échantillon sera donc appliquée à ce jeu B sous la forme d'un second cycle de modélisation.

Le deuxième cycle de modélisation se traduit en trois étapes :

À l'issue du premier cycle de modélisation, le schéma d'expérience correspondant au jeu B est déjà identifié au niveau pseudo-concret. Ce schéma a été mis en relation avec l'expérience aléatoire sous-jacente. Celle-ci a été intégrée en tant que telle dans la dimension discursive au référentiel théorique de l'élève, avec son schème de test associé construit dans la dimension instrumentale, ainsi qu'un mode de représentation par des signes appropriés résumant sa réalisation et le résultat associé à ce test dans la dimension sémiotique. La confrontation de cette réponse apportée à l'issue du premier cycle de modélisation, avec le schéma d'expérience initial mène à prolonger le questionnement portant sur cet indéterministe qu'est l'expérience aléatoire. Cela conduit à l'idée d'appliquer une approche scientifique basée sur un principe d'induction. Ce principe pilote la mise en œuvre d'une nouvelle étape de mathématisation. L'étape de mathématisation est mise en œuvre dans le domaine des statistiques descriptives. Un protocole de collecte des données correspondantes à la répétition de l'expérience aléatoire est défini. Il s'agit alors de construire un échantillon. Le schéma d'expérience correspondant à l'expérience aléatoire passe temporairement du statut d'objet d'étude au statut d'outil permettant cette construction. L'élève effectue enfin une synthèse des données qui en ressortent. Des outils du référentiel statistique peuvent alors être mobilisés pour résumer les données obtenues (nombre de succès, fréquence des succès, …), et le résultat qui en ressort mobilise des signes issus d'un registre de représentation spécifique. D'un point de vue algorithmique, nous pouvons aussi interpréter cette étape comme une boucle à nombre d'itérations défini, avec calcul de cumul du nombre de succès.

Les réponses apportées par ce modèle mathématique doivent être interprétées en regard du modèle pseudo-concret lié au schéma d'expérience. Les élèves sont amenés à rechercher une cohérence entre le modèle probabiliste envisagé et les résultats du schéma d'expérience testé. Mais à ce stade de l'activité, le résultat mathématique résumant un échantillon n'apporte pas de réponse claire. L'idée de tester un certain nombre d'échantillons peut alors émerger à cette étape du processus de modélisation.

Il s'agit de reprendre la démarche scientifique précédente mais en l'appliquant à l'échantillon en tant qu'objet d'étude. Cela relance un troisième cycle de modélisation qui mobilise à nouveau le domaine des statistiques descriptives pour produire et analyser une suite d'échantillons.

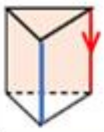
| Phases | ESPACE DE TRAVAIL MATHEMATIQUE | | Outil mobilisé | Objet en construction |
|---|---|---|---|---|
| Etape 3 | Plan cognitif | Idée de répéter l'expérience (simulation de l'expérience aléatoire). | Technique (ou instruction machine) de simulation de l'expérience aléatoire. | Construction progressive d'un échantillon de résultats issus de la simulation répétée de l'expérience aléatoire. |
| | Circulation du travail mathématique et domaine en jeu |  INS ↑ (Genèse instrumentale) Domaines : PROBA+STAT(+ALGO) | | |
| | Plan épistémologique | Assimilation du protocole de simulation de l'expérience aléatoire en tant qu'artefact. | | |
| Etape 4 | Plan cognitif | Nécessité ressentie de mémoriser les résultats successifs et de les interpréter. | Techniques classiques d'organisation de données basées sur une condition à vérifier (ou mobilisation éléments de syntaxe pour un programme informatique) | Présentation d'un archivage des données issues des simulations successives de l'expérience aléatoire (ou recherche d'un organigramme de programme). |
| | Circulation du travail mathématique et domaine en jeu |  INS – SEM ↓ (Representamen) Domaines : STAT (+ ALGO) | | |
| | Plan épistémologique | Mobilisation de signes pour mémoriser les données : traits, tableau, chiffres,... (ou éléments de syntaxe + schéma mental pour programme de simulation). | | |
| Etape 5 | Plan cognitif | Choix ou reconnaissance d'un critère d'arrêt pour la collecte répétée des données. Besoin de synthèse (résumer les résultats). | Comptage, comparaison, somme, fréquence (ou organisation d'instructions d'itération, condition, comparaison, boucle, cumul) | Finalisation de la construction d'un échantillon avec son bilan (ou finalisation de la rédaction du programme de simulation associé, en genèse discursive). Présentation des données. |
| | Circulation du travail mathématique et domaine en jeu |  SEM – DIS ↓ (↑) (Référentiel) Domaines : STAT (+ ALGO) | | |
| | Plan épistémologique | Mobilisation de principes de comparaisons, comptage, somme, fréquence (ou du savoir-faire algorithmique). | | |

*Figure 3. Description de l'Espace de Travail Mathématique correspondant au deuxième cycle de modélisation*

L'ETM statistique mis en œuvre lors de l'étape de mathématisation est a priori piloté par la dimension instrumentale qui entre en genèse : le principe de répéter un certain nombre de fois l'expérience aléatoire va automatiser des schèmes instrumentaux permettant la construction d'un échantillon. La nécessité d'employer des signes permettant la mémorisation, l'organisation, et le traitement des données mobilise alors la dimension sémiotique. Le principe algorithmique de boucle définie qui ressort de ce travail instrumental pourra ultérieurement intégrer le plan épistémologique de l'ETM statistique en tant qu'artefact. La nécessité de confronter le résultat obtenu avec le schéma d'expérience initial associé à l'expérience aléatoire correspondant au jeu B, entraine l'idée de résumer les données ressortant de l'échantillon obtenu. Cela va mobiliser des éléments du référentiel théorique tels que la somme, la moyenne, la fréquence…

## 3 Organiser et représenter les données issues de plusieurs échantillons (séance 4 et séance 5)

Le troisième cycle de modélisation se traduit en trois étapes :

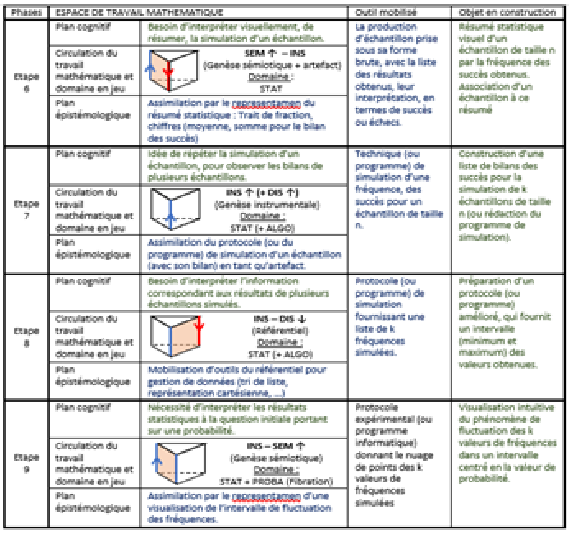
A l'issue du second cycle de modélisation, le schéma d'expérience correspondant au jeu B a permis de produire un échantillon qui doit être interprété au niveau pseudo-concret. Cela mène à rechercher une éventuelle propriété liée à cette production d'échantillon. Pour cela l'élève va a priori devoir observer et comparer plusieurs échantillons. Le protocole permettant la construction d'un échantillon a déjà été intégré en tant qu'artefact dans la dimension instrumentale, en tant qu'algorithme procédural embarquant ses schèmes, ainsi qu'un mode de représentation par des signes appropriés résumant sa réalisation et le résultat associé à cet échantillon dans la dimension sémiotique.

Une nouvelle étape de mathématisation est mise en œuvre dans le domaine des statistiques descriptives. Un protocole de production de plusieurs échantillons doit être mis en place. Le cadre de l'approche scientifique, par induction, de la notion d'échantillon, mène selon notre méthodologie à reproduire plusieurs échantillons de même taille, et sous les mêmes conditions, pour pouvoir les comparer entre eux. La procédure correspondant à la construction d'un échantillon passe alors du statut d'objet d'étude au statut d'outil permettant d'obtenir une série d'échantillons. L'élève doit effectuer une synthèse de ces données qui peut être élaborée dans différents registres de représentations (tableau, graphique en nuage de points, …). Des outils du référentiel statistique peuvent alors être mobilisés pour résumer les données obtenues (abscisse, représentation cartésienne,…). D'un point de vue algorithmique, nous pouvons aussi interpréter cette étape comme une boucle à nombre d'itérations défini, répétant la procédure de construction d'un échantillon.

Les réponses apportées par ce modèle mathématique doivent être interprétées en regard du modèle pseudo-concret lié au schéma d'expérience. Les résultats obtenus montrent une fluctuation des valeurs des effectifs (ou des fréquences) dans un certain intervalle à identifier et interpréter en regard d'hypothèses initiales pouvant porter sur l'expérience aléatoire considérée d'un point de vue probabiliste. Les élèves sont amenés à rechercher une cohérence entre le modèle probabiliste envisagé et l'intervalle de fluctuation d'échantillonnage. Il peut s'agir d'une observation intuitive en lien avec la notion de probabilité.

A l'issue de ce troisième cycle de modélisation, plusieurs approches sont envisageables. Il peut être intéressant d'exploiter une simulation informatique, pour obtenir davantage d'échantillons de plus grande taille, afin de densifier visuellement l'intervalle de fluctuation. A défaut d'un usage de l'informatique, il est possible d'effectuer des regroupements des résultats obtenus en classe par les différents élèves, ou binômes de travail, afin de dégager quelques échantillons de grande taille. Enfin, il est possible de passer dans le domaine probabiliste, en considérant un ensemble de chances équiprobables prenant en compte les résultats de chacun des deux dés, en les distinguant.

Ce qui permet d'identifier avec un univers à 36 issues 10 chances sur 36 de gagner au jeu B. Cela permet ensuite une comparaison avec les 2 chances sur 6 du jeu A. Une telle approche constitue une entrée dans le domaine probabiliste, et sa nature indéterministe, en tant que première rencontre. Le raisonnement associé s'appuie toutefois, sur une démarche de dénombrement simple.



| Phases | ESPACE DE TRAVAIL MATHEMATIQUE | | Outil mobilisé | Objet en construction |
|---|---|---|---|---|
| Etape 6 | Plan cognitif | Besoin d'interpréter visuellement, de résumer, la simulation d'un échantillon. | La production d'échantillon prise sous sa forme brute, avec la liste des résultats obtenus, leur interprétation, en termes de succès ou échecs. | Résumé statistique visuel d'un échantillon de taille n par la fréquence des succès obtenus. Association d'un échantillon à ce résumé |
| | Circulation du travail mathématique et domaine en jeu | SEM ↑ – INS (Genèse sémiotique + artefact) Domaine : STAT | | |
| | Plan épistémologique | Assimilation par le representamen du résumé statistique : Trait de fraction, chiffres (moyenne, somme pour le bilan des succès) | | |
| Etape 7 | Plan cognitif | Idée de répéter la simulation d'un échantillon, pour observer les bilans de plusieurs échantillons. | Technique (ou programme) de simulation d'une fréquence, des succès pour un échantillon de taille n. | Construction d'une liste de bilans des succès pour la simulation de k échantillons de taille n (ou rédaction du programme de simulation). |
| | Circulation du travail mathématique et domaine en jeu | INS ↑ (+ DIS ↑) (Genèse instrumentale) Domaine : STAT (+ ALGO) | | |
| | Plan épistémologique | Assimilation du protocole (ou du programme) de simulation d'un échantillon (avec son bilan) en tant qu'artefact. | | |
| Etape 8 | Plan cognitif | Besoin d'interpréter l'information correspondant aux résultats de plusieurs échantillons simulés. | Protocole (ou programme) de simulation fournissant une liste de k fréquences simulées. | Préparation d'un protocole (ou programme) amélioré, qui fournit un intervalle (minimum et maximum) des valeurs obtenues. |
| | Circulation du travail mathématique et domaine en jeu | INS – DIS ↓ (Référentiel) Domaine : STAT (+ ALGO) | | |
| | Plan épistémologique | Mobilisation d'outils du référentiel pour gestion de données (tri de liste, représentation cartésienne, …) | | |
| Etape 9 | Plan cognitif | Nécessité d'interpréter les résultats statistiques à la question initiale portant sur une probabilité. | Protocole expérimental (ou programme informatique) donnant le nuage de points des k valeurs de fréquences simulées | Visualisation intuitive du phénomène de fluctuation des k valeurs de fréquences dans un intervalle centré en la valeur de probabilité. |
| | Circulation du travail mathématique et domaine en jeu | INS – SEM ↑ (Genèse sémiotique) Domaine : STAT + PROBA (Fibration) | | |
| | Plan épistémologique | Assimilation par le representamen d'une visualisation de l'intervalle de fluctuation des fréquences. | | |

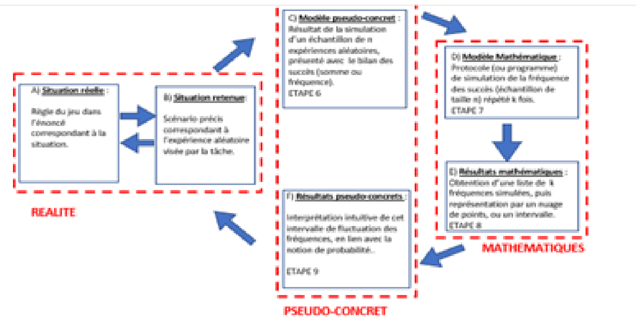



*Figure 4. Description de l'Espace de Travail Mathématique correspondant au troisième cycle de modélisation*

L'ETM statistique mis en œuvre lors de l'étape de mathématisation est a priori piloté par la dimension instrumentale qui entre en genèse : le principe de répéter un certain nombre de fois la construction d'un échantillon va automatiser des schèmes instrumentaux correspondant. La nécessité d'employer des signes permettant la mémorisation, l'organisation, et le traitement des données mobilise alors la dimension sémiotique dans différents registres (tableau, graphique). Le principe algorithmique de boucle double définie qui ressort de ce travail instrumental pourra ultérieurement intégrer le plan épistémologique de l'ETM statistique en tant que nouvel artefact de construction d'un nuage de points, et pourra même intégrer le référentiel théorique.

La nécessité de confronter le résultat obtenu avec le schéma d'expérience initial associé à l'expérience aléatoire correspondant au jeu B, apporte dans la dimension discursive de l'ETM statistique une preuve empirique basée sur le travail de visualisation qui ressort d'une genèse obtenue dans le dimension sémiotique.

# III - ANALYSE A POSTERIORI DU TRAVAIL DES ÉLÈVES

## 1 Parcours d'expérimentation initial en 2019/20 (CM2)

### 1.1 *Séance 1 (21/11/2019) : Découverte de deux jeux de hasard.*

Les élèves ont proposé dans un premier temps des réponses sans aucune justification. C'est la diversité de ces réponses qui a permis de lancer l'idée d'expérimenter. Certains élèves déplaçaient le débat vers une prise en compte de «chances» de gagner. Certains élèves annonçaient d'emblée qu'on ne peut répondre sur le hasard, et résistaient par rapport à la consigne. Une fois au travail pour la collecte des résultats, les élèves comptaient à la fois les succès et les échecs. Certains proposaient de limiter le nombre d'essais pour conclure plus facilement. Les élèves se basaient presque tous sur une petite série de lancers pour conclure.

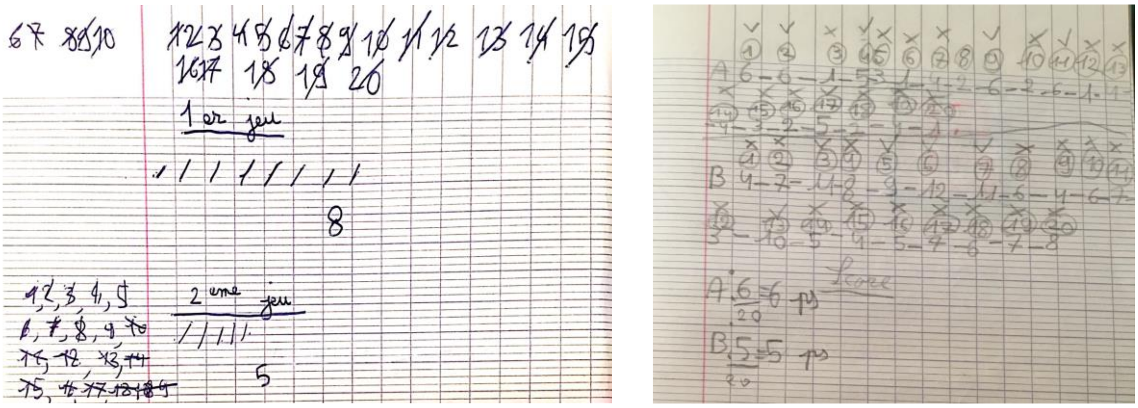
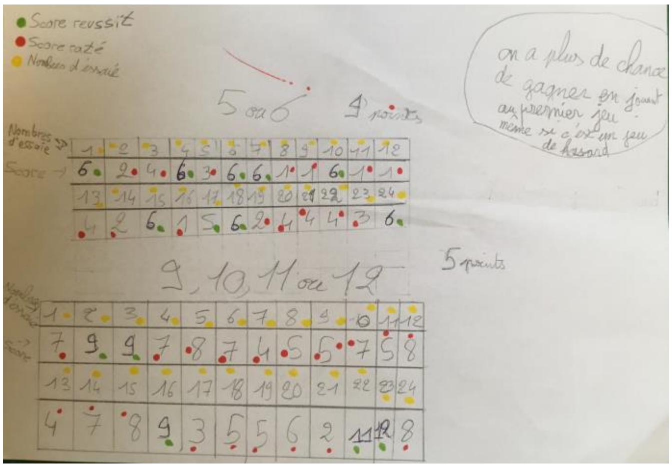



*Figure 5. Trace écrite de deux binômes différents lors de la séance 1 du 21/11/2019 (EPIAD)*

Concernant le processus de modélisation, les élèves sont entrés dans un modèle pseudo-concret en restreignant leur étude à des échantillons de taille 20 pour les deux jeux testés. Mais ils sont restés accrochés au monde réel du point de vue d'une prise en compte conjointe des deux jeux, et aussi en gardant un décompte à la fois des succès et des échecs. Concernant le travail mathématique, le registre sémiotique des chiffres a été exploité conjointement à des signes iconiques permettant d'identifier ou distinguer les différents statuts des nombres mentionnés, mais cela ralentit les opérations de traitements. Une genèse instrumentale est mise en œuvre pour la collecte des résultats correspondants aux tests effectués par rapport à l'expérience aléatoire. La dimension discursive est très peu mobilisée d'un point de vue mathématique, de manière restreinte au registre langagier.

### 1.2 *Séance 2 (09/12/2019) : Comparaison à partir de petits échantillons standards*

Les formes de tableaux sont variables ainsi que la mise en perspective de ces deux tableaux. Le codage des succès varie d'un tableau à l'autre. Les bilans apparaissent spontanément, et souvent sous forme de fractions. Les élèves sont perturbés par le fait que certains autres binômes n'ont pas le même résultat qu'eux (concernant le jeu apparaissant le plus avantageux). C'est ce point spécifique qui a convaincu la plupart des élèves de réaliser plusieurs tableaux de 20 lancers d'un même jeu, pour en savoir plus…



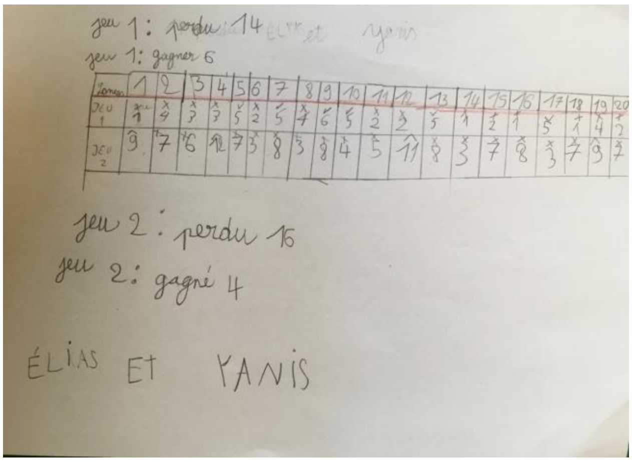




*Figure 6. Traces écrites de deux binômes différents lors de la séance 2 du 09/12/2019 (EPIAD)*

Une deuxième séance portant sur la comparaison d'échantillons de taille 20, permet de consolider les dimension instrumentale et sémiotique, ainsi que le modèle mathématique mobilisé dans le domaine des statistiques descriptives, en faisant appel au registre de représentation des tableaux, ce qui accélère les traitements.

### 1.3 *Séance 3 (23/01/2020) : Réalisation par binôme de 5 échantillons de taille 20 (jeu B)*

La séance 3 organisait la réalisation pour chaque binôme de 5 échantillons standardisés de taille 20, à renseigner dans des tableaux pré-remplis. L'objectif était de permettre ultérieurement des regroupements des résultats de chaque binôme pour former 5 échantillons de grande taille. Les binômes devaient aussi représenter dans un graphique les résultats obtenus. La conversion du registre des tableaux vers celui du graphique a posé certaines difficultés aux élèves, tout particulièrement en ce qui concerne les questions d'échelles.

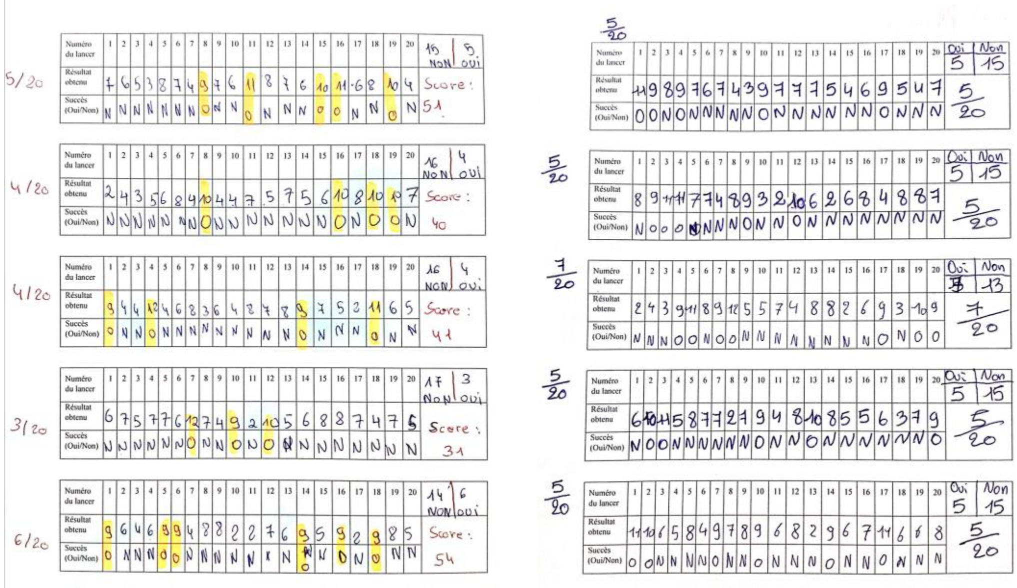


*Figure 7. Traces écrites de deux binômes différents lors de la séance 3 du 23/01/2020 (EPIAD)*

| | Nombre de succès | | | | | | | | | | | | TOTAL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | G10 | G11 | G12 | |
| Tableau 1 | 5 | 9 | 5 | 4 | 9 | 5 | 3 | 4 | 5 | 5 | 7 | | 61 |
| Tableau 2 | 7 | 11 | 5 | 2 | 5 | 4 | 1 | 8 | 7 | 8 | 6 | | 64 |
| Tableau 3 | 4 | 11 | 7 | 6 | 6 | 4 | 7 | 4 | 11 | 5 | 7 | | 72 |
| Tableau 4 | 2 | 9 | 5 | 10 | 4 | 3 | 8 | 4 | 7 | 3 | 6 | | 61 |
| Tableau 5 | 4 | 10 | 5 | 6 | 9 | 6 | 8 | 8 | 4 | 5 | 4 | | 69 |

*Figure 8. Bilan des résultats obtenus pour chaque binôme lors de la séance 3 du 23/01/2020 (EPIAD)*

L'idée de mettre en commun les données a mis du temps à émerger du côté des élèves. Il a fallu insister à l'appui du tableau collectif. Ces premières réflexions portant sur une exploitation du tableau commun ont nécessité un arbitrage entre des sommes par colonnes et des sommes par lignes. C'est une des raisons pour lesquelles il a été envisagé de reconduire une variante de cette expérimentation l'année suivante, en ajoutant une séance menant à faire au préalable travailler les élèves sur une synthèse de tous les échantillons obtenus.

### 1.4 *Séance 4 (30/01/2020) : Représentation graphique des données obtenues (jeu B)*

Les représentations graphiques libres ont donné lieu à des représentations très diverses (diagrammes en bâtons, traits de constructions verticaux et horizontaux, points reliés par des segments, …) dont la visualisation ne met pas toujours bien en évidence l'aspect fluctuation. La représentation graphique standard par des points, telle qu'adoptée par la classe après les représentations libres, amène encore les élèves à relier spontanément ces points. Mais beaucoup de binômes relient aussi l'origine du repère. Nous remarquons que les élèves ont des réticences à déconstruire dimensionnellement les représentations spatiales (du rectangle au trait, du trait au point). Lors de cette séance, les élèves réalisent que les nombres de succès de chaque binôme n'a jamais atteint 12.

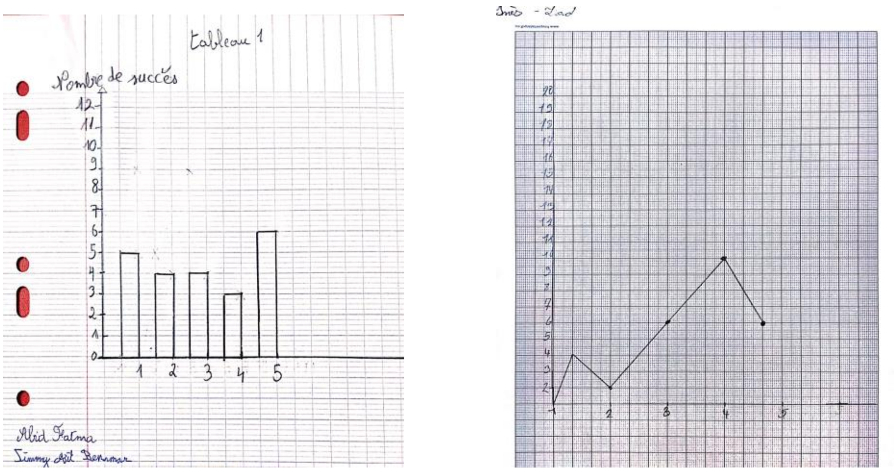



*Figure 9. Les représentations graphiques pour deux binômes lors de la séance 3 du 23/01/2020 (EPIAD)*

## 2 Reprise de parcours d'expérimentations en 2020/21 (6ème , CM2 , CM1)

### 2.1 *Reprise, prolongement, et compléments de l'expérience au niveau CM2*

L'objectif de ce nouveau parcours d'expérimentation en CM2 a été de confirmer certaines observations réalisées en 2019/20, et aussi de prolonger les travaux jusqu'à la réduction des fluctuations par augmentation de la taille des échantillons, tel que cela avait initialement été prévu.

Un soin particulier a été apporté à une séance intermédiaire permettant la transition des petits vers les grands échantillons. Sur suggestion d'un binôme, l'enseignant a regroupé tous les nuages des points correspondant aux petits échantillons dans un graphique commun, ce qui permet de localiser spatialement le nuage.

*Figure 10. Les représentations graphiques pour deux binômes lors de la séance 5 PEH)*

Ce travail de synthèse portant sur 50 petits échantillons de taille 20 a montré la richesse des procédures pouvant être employées par les élèves : ligne polygonale reliant les points, recherche d'une position médiane, identification de bornes élaguées, diagramme en bâton des effectifs (nombre d'échantillons) correspondant à chaque nombre de succès … Cette dernière démarche a été une surprise du fait de la conversion de registre sous-jacente. L'usage à bon escient de signes (marques, couleurs) par les élèves, en lien avec les règles de traitement du registre sémiotique du graphique, a permis le lancement de routines instrumentales diverses et variées, exécutées sur le papier.

L'activité a aussi entrainé des échanges importants entre élèves, qui participaient du travail discursif de validation. L'expérimentation de la séance 5 a aussi confirmé une observation menée au secondaire : les élèves engagés dans un traitement statistique avancé tel que la recherche de bornes élaguées perdait de vue le modèle pseudo-concret faisant référence à une situation de jeu menant à des succès et des échecs. Ils se retrouvaient confinés dans le domaine des statistiques descriptives.

Un enjeu important semble se situer au niveau de cette fibration sémiotique mettant en correspondance les bornes du nuage statistique avec une forme d'estimation probabiliste. L'idée de regrouper les échantillons pour former quelques échantillons de grande taille a quant à elle émergé assez tardivement au cours de cette même séance.

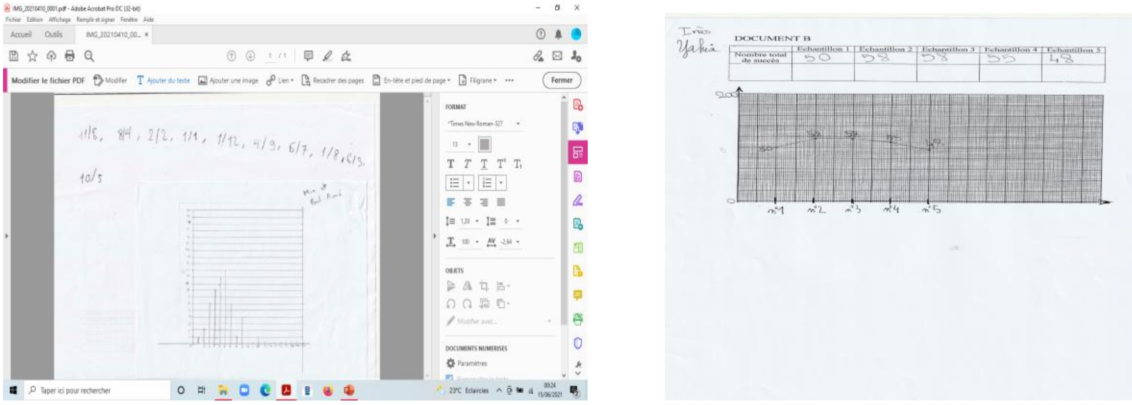



*Figure 11. Les représentations graphiques pour deux binômes de la PEH lors des séances 5 (à gauche) et 6 (à droite).*

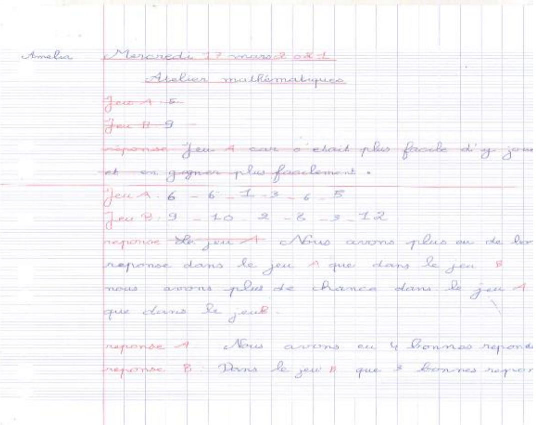
La séance 6 a permis de représenter les fluctuations correspondantes au regroupement en 5 échantillons de taille 200. L'objectif pédagogique de cette séance correspondait au choix de l'échelle en ordonnée. Cette difficulté entraine des erreurs et des maladresses liées aux graduations du papier millimétré. La relation entre la partie et le tout n'est pas toujours mise en évidence en ordonnée. L'échelle choisie n'est pas forcément respectée au moment de représenter les points. Tout cela perturbe l'interprétation attendue du phénomène probabiliste.

### 2.2 *Compléments d'observation menés au niveau 6ème*

L'expérimentation menée en parallèle au niveau d'une classe de 6ème a été prolongée jusqu'au bilan d'étape ayant permis de réunir lors d'une visioconférence les trois classes engagées dans ce projet pédagogique. La réalisation des échantillons de taille 20 a donné lieu à une exploitation plus directe et rapide de la notion de fréquence. Par ailleurs l'interprétation probabiliste s'appuyant sur une prise en compte de chacun des résultats affichés par les deux dés, pour comptabiliser 10 cas favorables sur 36 cas possibles, a été formulée par plusieurs binômes.

### 2.3 *Compléments d'observation menés au niveau CM1*

L'expérimentation menée en parallèle au niveau d'une classe de CM1 a permis de mettre en évidence des différences importantes du point de vue du travail mathématique mis en œuvre par les élèves. Une hypothèse est que sur le plan cognitif les sujets sont encore à ce stade de leur développement psychologique très ancrés dans le stade des opérations concrètes. Cela se traduit par une difficulté à focaliser sur un modèle pseudo-concret suffisamment clarifié. De plus, la visualisation des signes employés, et tout particulièrement des chiffres, a tendance à confiner les élèves dans des considérations de calcul liées à ces mêmes chiffres. C'est-à-dire qu'au lieu de parvenir à dénombrer les succès en passant en revue les sommes obtenues pour les valeurs affichées sur les dés, les élèves ont tendance à cumuler ces mêmes sommes, puis à les comparer entre les deux jeux, ce qui relève d'une perte de sens en regard de la situation réelle. Enfin l'idée légitime d'imprévisibilité du hasard semble rester, au niveau du CM1, un obstacle épistémologique très difficile à surmonter. Les approches restent liées à ce questionnement purement déterministe de la part des élèves.



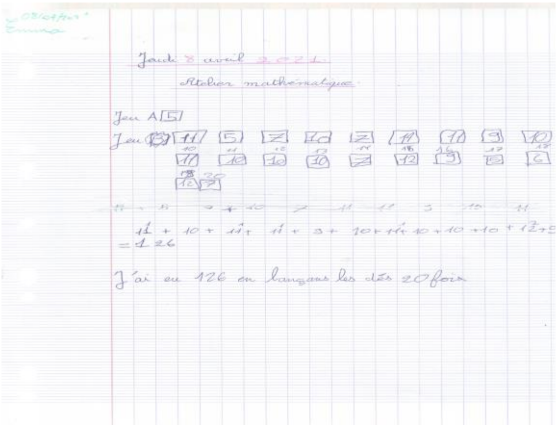




*Figure 12. Trace écrite de deux binômes de CM1 à l'EPIAD lors de la séance 1 (à gauche) et de la séance 2 (à droite).*

## 3 Un bilan provisoire pour l'analyse didactique des résultats

Les travaux expérimentaux menés au niveau du lycée en lien avec un travail de thèse portant sur les approches fréquentiste et combinatoire de la notion de probabilité, avaient permis de constater une structuration particulière en trois étapes, liée à la réalisation en classe d'une simulation d'échantillonnage.

Des embryons de connaissance intuitive sont progressivement mobilisés pour enchaîner des techniques, et des étapes de construction du protocole (ou du programme de simulation) par encapsulages successifs. Ces étapes de nature procédurales entrent en résonance avec les principales phases du protocole expérimental statistique.

Ces travaux expérimentaux complémentaires menés au niveau de cycle 3 ont permis d'élargir la vision des difficultés rencontrées sur le plan cognitif lorsqu'on aborde ces notions. Certaines conceptions sous-jacentes apparaissent en effet occultées par de nombreuses notions déjà acquises, au secondaire, en lien avec l'automatisation d'expertises liées aux domaines du calcul et de la gestion de données. Expérimenter au niveau du cycle 3 a permis d'identifier les enjeux associés à l'interaction entre différents domaines mathématiques qui sont encore en cours de développement ou de consolidation tels que les nombres, les opérations, l'idée de cumul, de classement, le calcul numérique, la notion de fraction, de quotient, les conventions du type tableau, la représentation graphique cartésienne…

L'outil ETM met en évidence, dans ce cadre, les différentes dimensions du travail mathématique. Les étapes de structuration du protocole de simulation sont tour à tour considérées comme objet d'étude soumis à des traitements et des conversions sémiotiques, et artefact (outil) permettant la genèse instrumentale d'un nouveau protocole expérimental.

L'identification de cette dialectique outil-objet permet de décomposer l'activité en autant de cycles de modélisation. Après validation, l'ancien protocole devient un nouvel outil qui va entrer en genèse discursive dans l'ETM pour produire un nouvel objet procédural permettant d'automatiser ce nouveau protocole expérimental… L'ETM met alors en évidence une triple répétition de ce processus : une dialectique cyclique d'interactions évolutives, qui correspond aussi à l'enchaînement de trois cycles de modélisations : le modèle de l'expérience aléatoire, le modèle de l'échantillon et de son résumé, le modèle numérique ou graphique permettant d'observer la fluctuation des fréquences (ou nombre de succès) associée à un ensemble d'échantillons.

Du point de vue de la recherche en didactique, ces travaux ont donné lieu à de nombreuses données dont l'analyse complète prendra encore du temps. Ce chantier pourrait prendre toute sa dimension à l'issue de la thèse en cours, afin de compléter la réflexion portant sur l'introduction de la probabilité au secondaire, par un lien avec l'enseignement de la gestion de données au primaire.

# IV - BILAN DE LA FORMATION DE FORMATEURS

Les travaux se sont finalement appuyés à la fois sur une validation institutionnelle (conseil école-collège) et sur des travaux similaires en cours au niveau du secondaire (en lien avec des expérimentations de thèse). Plusieurs professeurs référents expérimentateurs ont été identifiés dans chacun des trois établissements concernés. Les expérimentations ont été menées en 2 vagues : phase exploratoire en 2019-20, et parcours d'étude et de recherche en 2020-21. Le déroulement du projet a été ponctué de 4 bilans d'étapes réalisés en présence d'universitaires dans le cadre des séminaires périodiques du Laboratoire de Mathématiques Audin. Un lien a été mis en place avec le dispositif de formation, dans l'objectif de confectionner un outil pédagogique exploitable lors de stages inter-degrés.

D'un point de vue plus personnel lié aux impressions rendues par les professeurs expérimentateurs ayant été engagés dans ce projet, l'atelier A15 de la COPIRELEM 2021 a permis différents témoignages liés au bénéfice apporté par ce projet du point de vue du renforcement des compétences professionnelles.

Un enseignant avait remarqué que ses élèves ne s'engageaient pas, d'ordinaire dans la résolution de problèmes, car il n'y avait pas vraiment de dynamique créée par les exercices proposés. L'expérimentation décrite dans cette communication lui a permis d'observer que la proposition de problèmes résistant aux élèves, et pour lesquels aucune procédure experte de résolution n'est disponible, les incitaient à se "lancer" dans la recherche d'une stratégie de résolution. Selon cet enseignant, le problème présenté sous la forme d'un défi à surmonter avait motivé les élèves, avec une réflexion de leur part qui se poursuivait jusque dans la cour de récréation.

Ce projet pédagogique a amené les enseignants expérimentateurs à proposer davantage de problèmes ouverts, pour lesquels il n'y a pas de solution immédiate, mais plutôt plusieurs résolutions possibles. Le cadre de tels problèmes se résolvant par essais, tâtonnements, erreurs, échecs, permet en effet de prendre conscience que l'erreur est une étape nécessaire dans l'apprentissage. De plus, il est aussi apparu dans les remarques des enseignants expérimentateurs que l'énoncé des problèmes est primordial : cet énoncé doit piquer la curiosité de l'élève, susciter chez lui un sentiment de défi intellectuel et surtout ne pas induire la méthode de résolution.

Les enseignants engagés dans cette expérimentation ont revu la place des problèmes dans l'apprentissage des mathématiques, en prenant conscience que la résolution de problèmes doit passer par un apprentissage spécifique, ce qui est central dans l'apprentissage des mathématiques.

Un collègue expérimentateur explique que dans sa pratique antérieure, la notion de recherche n'était pas toujours présente : de nombreux élèves considéraient qu'on devait pouvoir répondre immédiatement à une question. A présent, ce collègue insiste en lien avec l'idée de compétences, sur le fait que résoudre un problème c'est en effet chercher la solution, mais aussi et surtout expliquer comment on a procédé pour convaincre lorsque cette réponse n'est ni immédiate ni évidente.

Outre ces témoignages issus de l'expérience de ce projet pédagogique, nous rappelons en effet que ce dispositif ciblait un objectif de formation de formateurs. Il apparaît aussi que la recherche en didactique des mathématiques a montré une double utilité dans le cadre de cette ingénierie : en permettant de mettre des mots sur certains concepts pédagogiques, en poussant la réflexion collégiale entre pairs, et aussi par la complémentarité des rôles respectifs du chercheur et de l'enseignant expérimentateur, entraînant une coopération de fond dans des pratiques professionnelles qui sont respectivement valorisées l'une en regard de l'autre. Pour l'anecdote, en plus de l'action de formation sous-jacente à ce dispositif, ayant bénéficié à ces quatre collègues expérimentateurs, deux d'entre eux sont allés plus loin depuis cette expérience. L'un professeur des écoles est devenu Référent Mathématique de Circonscription dans l'Académie de Lille. L'enseignante ayant mené l'expérimentation en 6ème s'est engagée au sein de l'axe didactique du Laboratoire Audin, dans le parcours Master 2 en didactique des mathématiques encadré à distance par l'Université d'Aix-Marseille.

Les échanges avec les membres de l'atelier A15 n'ayant pas participé en amont au dispositif ont donné lieu à plusieurs éclaircissements liés à la mise en place au niveau du primaire d'une telle expérimentation. Une première question se posait en lien avec le risque d'une mise en correspondance des échantillons avec une notion de fraction : en effet réunir deux échantillons ne correspond pas à l'addition des fractions exprimant les fréquences de succès obtenus pour chacun de ces échantillons… Une autre question s'est posée en lien avec la problématique du caractère indiscernable des deux dés. L'idée d'utiliser deux dés de couleurs différentes, ou encore deux dés différents en termes de nombres de faces, est alors apparue pertinente… Mais l'expérience s'intéressait aussi au type de confusion qui peut ressortir de la situation associée à deux dés indiscernables chez des élèves de primaire.

Une question d'ordre didactique s'est posée par rapport au cadre théorique choisi : il est alors devenu nécessaire de préciser pourquoi le processus de modélisation a été décomposé en trois cycles (test de l'expérience aléatoire, formation de l'échantillon, formation d'une liste de fréquences des succès). Enfin une autre question s'est posée par rapport aux causes de cette différence significative entre les élèves de CM2 et les élèves de CM1 pour une même activité posée, sachant que dans l'absolu les programmes de ces deux classes ne diffèrent pas vraiment par rapport aux outils mathématiques en jeu. L'explication proposée s'est alors plutôt appuyée sur le plan cognitif : les élèves de CM1 apparaissent en effet beaucoup plus soumis aux chiffres qui sont visualisés physiquement, tels que les sommes écrites des deux dés, qui occultent alors le dénombrement sous-jacent des succès obtenus.

## V - BIBLIOGRAPHIE


Henry, M. (1999). L'introduction des probabilités au lycée : un processus de modélisation comparable à celui de la géométrie. *Repères-IREM*, *36*, 15–34.

Henry, M. (2010). Évolution de l'enseignement secondaire français en statistique et probabilités, contributions des IREM à la formation continue des enseignants. In *Statistique et Enseignement*, Vol.1 (pp. 35–45). Société Française de Statistique.

Kuzniak, A. (2011). L'espace de Travail Mathématique et ses genèses. *Annales de didactique et de sciences cognitives*, vol. 16, 9-24.

Kuzniak, A., & Richard, P. R. (2014). Espaces de Travail Mathématique. Points de vue et perspectives. *Revista Latinoamericana de Investigación en Matemática Educativa, Especial 2* (Tome I), 29–40.

Kuzniak, A., Nechache, A., & Drouhard, J. P. (2016). Understanding the development of mathematical work in the context of the classroom. *ZDM – Mathématics Education*, *48* (6), 861–874.

Nechache, A. (2016). La validation dans l'enseignement des probabilités au niveau secondaire. *Thèse de doctorat de l'université Paris Diderot*, Paris, France.

Parzysz, B. (2007). Expérience aléatoire et simulation : le jeu de croix ou pile. Relecture actuelle d'une expérimentation déjà un peu ancienne. *Repères-IREM*, *66*, 27–44.

Parzysz, B. 2009. De l'expérience aléatoire au modèle, *via* la simulation. *Repères- IREM*, 74, 91-103.

Parzysz B. (2011). Une générateur aléatoire de pile ou face venu d'ailleurs. *Bulletin de l'APMEP*, 494, 309-314.

Parzysz, B. (2011). Quelques questions didactiques de la statistique et des probabilités. *Annales de didactique et des sciences cognitives*, *16*, 127–147.

Trunkenwald, J., Laval, D. (2019). Algorithms as a discovery process in frequentist approach to prediction interval. In Jankvist, U. T., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (Eds.). (2019), *Eleventh Congress of the European Society for Research in Mathematics Education (CERME11)*, Utrecht University, Feb 2019, Utrecht, Netherlands. hal-02412851

Trunkenwald, J. (2018). La simulation d'échantillonnage : Une synergie entre les domaines probabiliste et statistique. *Actes du Sixième symposium international – Espace de Travail Mathématique* de décembre 2018, pp. 165-177. Valparaiso, Chili.

Trunkenwald, J. (2019). Entre probabilités et statistiques : Un jeu algorithmique pour simuler la fluctuation d'échantillonnage. In M. Abboud (Éd.), *Actes du colloque EMF 2018*, (pp.1681-1689). Paris : IREM de Paris.