

Table des matières

I. RESUME	3
II. MOTS-CLES	3
III. LISTE DES TABLEAUX	3
IV. LISTE DES FIGURES	4
V. LISTE DES ANNEXES	4
INTRODUCTION	5
1. PROBLEMATIQUE	6
1.1 POINT DE DEPART DE LA REFLEXION	6
1.2 DEFINITION ET IMPORTANCE DE L'OBJET DE RECHERCHE	8
1.3 ETAT DE LA QUESTION	11
1.3.1 <i>Le fonctionnement de la mémoire</i>	11
1.3.1.1 Types de mémoires	11
1.3.1.2 Consolidation et rappel des informations	12
1.3.1.3 Le mécanisme de l'oubli	13
1.3.1.4 Stratégies de mémorisation à long terme	14
1.3.2 <i>Le testing effect</i>	14
1.3.2.1 Un sujet d'étude centenaire	15
1.3.2.2 Le testing effect dans un environnement scolaire réel	16
1.3.2.3 Importance du feedback	18
1.3.2.4 Conclusions sur le testing effect	21
1.3.3 <i>Les quiz dans l'apprentissage des langues</i>	21
1.3.3.1 Présentation	22
1.3.3.2 Avantages des quiz	22
1.3.3.3 L'apport de l'informatique	24
1.3.3.4 Conception des quiz et définition des items	31
1.4 QUESTIONS ET OBJECTIFS DE RECHERCHE	36
2. PREMIERE PARTIE	38
2.1 METHODOLOGIE	38
2.1.1 <i>Fondements méthodologiques</i>	38
2.1.2 <i>Nature du corpus</i>	39
2.1.2.1 Choix de la branche d'enseignement	39
2.1.2.2 Population	39
2.1.2.3 Contenu didactique	39
2.1.2.4 Procédure et protocole de l'expérience	40
2.1.2.5 Différences par rapport à l'étude de base	42
2.1.3 <i>Méthode de collecte des données</i>	43
2.2 ANALYSE	44
2.2.1 <i>Ecarts de performance entre les 3 types de questions</i>	44

2.2.1.1 Analyse des moyennes en ligne pour chaque question	45
2.2.2 Incidence du dispositif sur la performance des élèves	48
2.2.3 Ecart de performance selon les catégories d'élèves	49
2.2.4 Discussion.....	50
3. DEUXIEME PARTIE	52
3.1 METHODOLOGIE.....	52
3.1.1 Fondements méthodologiques.....	52
3.1.2 Nature du corpus	52
3.1.2.1 Population.....	52
3.1.2.2 Procédure et protocole de l'expérience	52
3.1.2.3 Problématique de l'accès à Internet	53
3.1.2.4 Accès aux quiz.....	54
3.1.2.5 Disponibilité des téléphones portables des élèves	54
3.1.3 Méthode de collecte des données	55
3.2 ANALYSE	56
3.2.1 Résultats et changements apportés	56
3.2.2 Evolution des questions posées	63
3.2.3 Discussion.....	65
4. PISTES D'AMELIORATION DU DISPOSITIF.....	66
CONCLUSION.....	68
REFERENCES BIBLIOGRAPHIQUES	70
ANNEXES.....	73

i. Résumé

Sur la base d'une étude de Roediger III et al. (2011) démontrant l'existence d'un *testing effect* dans une classe de sciences sociales dans une *middle school* américaine, Mon mémoire avait pour objectif de voir si la rétention des contenus d'un cours de langue du secondaire 1 peut être renforcée sur le long terme par l'ajout de tests de rappels à un enseignement existant. Il comporte deux parties distinctes. La première cherchait à vérifier mon hypothèse que les résultats de Roediger III et al. (2011) s'appliquent à une classe de langue seconde du secondaire 1, à savoir que les quizz à réponses courtes obtiennent des résultats plus élevés comparés aux questionnaires à choix multiples, et que la matière qui n'a bénéficié d'aucun dispositif obtient le plus bas pourcentage de bonnes réponses lors d'une évaluation sommative. La seconde partie de mon étude visait à définir la forme à donner aux tests de rappel pour obtenir les meilleurs résultats, sous la forme d'une étude qualitative menée sous forme d'itérations. Au final, un *testing effect* a pu être observé, confirmant ainsi l'hypothèse. Le dispositif mis au point au fil des semaines (outil utilisé, interface, contenu proposé, fréquence des tests) est également décrit.

ii. Mots-clés

- test-enhanced learning
- rétention à long terme d'information (long-term retention)
- testing effect
- computer-based tests
- apprentissage des langues secondes

iii. Liste des tableaux

Tableau 1: absences de certains élèves aux quiz	44
Tableau 2: moyennes des résultats des élèves, pour les 3 types de questions	44
Tableau 3: résultats pour chaque type de question	45
Tableau 4: résumé du quiz 1	56
Tableau 5: résumé du quiz 2	58
Tableau 6: résumé du quiz 3	60
Tableau 7: résumé du quiz 4	61
Tableau 8: résumé du quiz 5	63

iv. Liste des figures

Figure 1 : image de couverture (source : http://failureconfetti.smackjeeves.com/).....	0
Figure 2: la courbe de l'oubli d'Ebbinghaus (https://www.penserchanger.com/marre-doublier-voici-la-courbe-debbinghaus).....	7
Figure 3: source: Metcalfe, Kornell & Finn (2010), p. 1082	19
Figure 4: source: Metcalfe, Kornell & Finn (2010), p. 1084	20
Figure 5: niveaux de complexité dans l'évaluation des apprentissages. Source : Baillat, De Ketele et al. (2008). Evaluer pour Former. DBS Pédagogie.....	32

v. Liste des annexes

Annexe 1 : quiz 1 au format PDF

Annexe 2 : quiz 2 au format PDF

Annexe 3 : quiz 3 au format PDF

Annexe 4 : quiz 4 au format PDF

Annexe 5 : évaluation sommative

Annexe 6 : résultats de la classe test sur toute la durée du dispositif

Annexe 7 : résultats des élèves pris en compte dans l'étude à l'évaluation sommative

Annexe 8 : calculs pour l'analyse de la première partie de l'étude

Annexe 9 : score des questions de l'évaluation finale, par ligne

Introduction

La majorité des situations d'utilisation d'une langue seconde nécessitent une compréhension et une capacité de réponse rapides, l'interlocuteur se trouvant à proximité spatiale immédiate, attendant une réponse de notre part. Il n'est bien souvent pas possible, ou en tout cas pas optimal, de consulter un dictionnaire ou de rapidement pianoter sur son smartphone à la recherche du terme désiré. Cela implique que les compétences langagières doivent être aisément mobilisables. Malheureusement, d'après mon expérience d'élève, puis d'enseignant, les élèves du secondaire 1 travaillent trop souvent en vue de la réussite des évaluations et non dans une perspective de mémorisation à plus long terme ; cela a comme conséquence un oubli rapide et massif des informations, en lieu et place d'une acquisition pérenne. Ainsi, mon questionnement initial était de trouver un dispositif permettant de renforcer la rétention à long terme d'informations langagières, notamment le vocabulaire, qui constitue la clé de voûte de toute communication (Seashore, 1948).

En ce sens, mon mémoire a pour objectif de voir si la rétention des contenus d'un cours de langue du secondaire 1 peut être renforcée sur le long terme par l'ajout de tests de rappels à un enseignement existant. De nombreuses études ont démontré les avantages de tester des apprenants sur la matière traitée en fin d'une séquence, puis de répéter ces tests à diverse fréquence pendant les jours et semaines suivantes. Ce processus, connu sous le nom de *testing effect*, permettrait d'augmenter significativement la rétention d'informations sur le long terme. Une étude de Roediger III et al. de 2011 démontrant l'existence d'un *testing effect* dans une classe de sciences sociales dans une *middle school* américaine – l'équivalent de notre école secondaire – constitue la pierre angulaire de ma recherche.

Mon étude se décompose en 2 parties ; la première cherche à vérifier si, de manière similaire à l'étude de Roediger III et al. (2011), un *testing effect* peut être mesuré dans une classe helvétique du secondaire 1, à l'aide de tests de rappels effectués sous la forme de quiz informatisés. Au contraire de l'étude de Roediger III et al. (2011) cependant, où l'expérience a été menée dans le cadre d'un cours de sciences sociales, mon étude a pour cadre un cours d'allemand.

La seconde partie consiste à proposer un dispositif de *test-enhanced learning* (enseignement amélioré par les tests) pour un cours d'allemand du secondaire 1, qui vienne s'inscrire dans une pratique quotidienne déjà en place, sans chamboulement coûteux en temps pour l'enseignant. Il s'agit là d'un processus itératif ayant pour but, sur la base des résultats observés et du retour des élèves testés, d'améliorer le dispositif de semaine en semaine. L'objectif est de définir quels outils choisir pour construire et passer ces tests, définir l'interface optimale, conceptualiser leur contenu (types de

questions proposées, manière de formuler les questions) ainsi que leur mode de passation (lieu, temps alloué).

Au niveau du cadre théorique, je vais d'abord m'intéresser au fonctionnement de la mémoire, en particulier de la mémoire à long terme. Les mécanismes de stockage de l'information et les facteurs qui influencent son oubli ou au contraire sa rétention seront passés en revue, avec une attention particulière sur les éléments qui agissent positivement sur la récupération d'informations en mémoire. Ensuite, les études préalables sur le *testing effect* seront présentées et commentées, notamment les études récentes effectuées au sein de véritables classes sur du matériel scolaire authentique, au contraire de celles menées en laboratoire ou sur du contenu artificiellement créé pour les tests. Enfin, la question des quiz dans l'enseignement des langues sera traitée ; les pratiques en vigueur seront présentées, ainsi que les avantages liés à l'utilisation de quiz en classe. Après quelques éclairages sur l'essor des quiz grâce à l'informatique, une sélection d'outils de création disponibles sur le marché seront présentés, ainsi que les études traitant de la conception des quiz, en particulier ceux comprenant des questions de type QCM et choix ouvert.

Ensuite, je décrirai la méthodologie de travail imaginée et déployée dans ma classe en argumentant chacun de mes choix. Enfin, je présenterai les résultats de cette étude ainsi que les concepts clés liés à la mise en place d'un environnement d'enseignement amélioré par les tests, dans l'optique de permettre à tout enseignant de langue qui le souhaiterait d'intégrer les tests de rappel dans sa pratique quotidienne.

1. Problématique

1.1 Point de départ de la réflexion

En tant qu'enseignants, l'un de nos objectifs est que notre enseignement soit pérenne dans le temps, c'est-à-dire que les notions enseignées ne soient pas oubliées après quelques jours ou semaines, mais ancrées de manière durable dans la mémoire des apprenants. D'un point de vue cognitif, cela implique que l'information à mémoriser passe de la mémoire à court terme, qui ne retient les informations que temporairement, à la mémoire à long terme, qui permet de se rappeler d'éléments de manière durable. Cette dernière se nomme « mémoire sémantique » : elle contient des informations sur nos connaissances générales, comme par exemple le nom d'un objet ou la signification d'un terme précis. Le passage d'un élément de la mémoire à court terme vers la mémoire à long terme est influencé par divers facteurs et peut prendre plus de temps selon les individus. Ce processus est connu sous le nom de « stockage » (Tacconat, 2012) ou « consolidation » ("Consolidation de la mémoire," n.d.) dans la

littérature sur le sujet. Le stockage sert à créer des liens entre les différentes informations en mémoire. C'est ce qui permet de mieux se remémorer les choses par la suite (Taconnat, 2012).

Au sein du système scolaire helvétique, l'apprenant est évalué dans chaque branche au cours de l'année scolaire et voit ses efforts rétribués par une note. Bien souvent, celle-ci est attribuée sur la base d'une évaluation sommative intervenant en fin d'un chapitre ou d'une unité. Son but est d'évaluer si la matière étudiée précédemment en cours est acquise. Avec ce système, les élèves ne travaillent pas dans le but d'améliorer leurs connaissances de manière durable mais ont pour principal objectif de réussir les évaluations sommatives proposées. Le métier d'élève se borne ainsi à « la capacité de savoir refaire, au bon moment, ce qui a été longuement exercé » (Perrenoud, 1994). En ce sens, leur motivation première est d'emmagasiner suffisamment d'informations pour réussir une épreuve, puis de « purger » leur esprit une fois l'évaluation passée, de manière à préparer l'épreuve suivante. Ce bachotage a comme avantage d'être bien plus économique qu'une révision standard ; on peut le faire à la dernière minute : « aussitôt appris, aussitôt récité, aussitôt oublié ! » (Perrenoud, 1994).

Cependant, même sans bachotage, les informations mémorisées s'effacent progressivement si elles ne sont pas réactivées régulièrement, selon une courbe de l'oubli conceptualisée au XIX^e siècle déjà par Ebbinghaus (Nicolas, 1992). Des rappels effectués de manière sporadique permettent un gain important et une meilleure rétention de l'information (Nicolas, 1992). Le schéma ci-dessous représente les deux courbes : la courbe normale de l'oubli et celle obtenue dans le cas de rappels périodiques. Après 6 mois, on peut constater une différence de l'ordre de 60% entre les deux courbes. Cette différence représente le gain obtenu grâce au rappel :

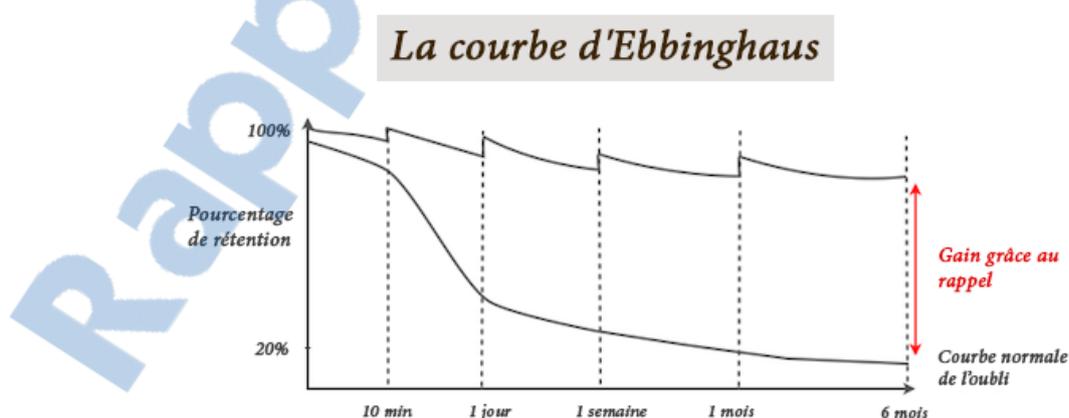


Figure 2: la courbe de l'oubli d'Ebbinghaus (<https://www.penserchanger.com/marre-doublier-voici-la-courbe-debbinghaus>)

Sachant cela, il est donc à mon sens utile de trouver une parade à cet oubli massif, par exemple en joignant le parcours des apprenants de rappels périodiques. Reste à définir sous quelle forme.

Afin que mon étude ne se cantonne pas à une expérience de laboratoire mais qu'un tel dispositif puisse être déployé pour d'autres enseignants, il est également important à mes yeux que les points suivants soient respectés :

- Le dispositif doit être peu intrusif et pouvoir s'intégrer à un enseignement existant.
- Afin d'éviter d'utiliser du papier supplémentaire, en plus des nombreuses fiches distribuées au long de l'année, le support numérique est à privilégier.
- Le dispositif doit pouvoir être mis en place par un enseignant sans connaissances techniques poussées, sans équipe technique associée.
- Le dispositif ne doit pas être chronophage pour l'enseignant, que ce soit pour sa préparation ou son application en classe.
- Le dispositif doit être peu coûteux et ne pas engendrer l'achat de matériel spécifique.

1.2 Définition et importance de l'objet de recherche

L'école obligatoire, bien que perçue comme un fardeau par certains de nos élèves, est une chance donnée aux jeunes d'acquérir un savoir important de manière gratuite – en faisant bien entendu abstraction des impôts payés par leurs parents. En ce qui concerne les langues secondes, les élèves ont accès à des cours d'allemand dès la 5^e année Harmos, à des cours d'anglais dès la 7^e année Harmos, ainsi que la possibilité de s'essayer à l'italien ou l'espagnol en 11^e Harmos. A la fin de leur parcours scolaire, ils bénéficieront au minimum d'un bagage langagier de 7 ans d'allemand et de 3 ans d'anglais. Cela représente cependant un coût important à la charge des contribuables. A titre d'exemple, les pouvoirs publics helvétiques ont investi 36 milliards de francs dans l'éducation en 2014, soit 5.6% du produit intérieur brut de la Suisse. De cette somme, 49% a été consacrée à l'école obligatoire, soit en moyenne 20'600 francs par élève¹.

En comparaison, suivre un cours de langue dans le privé représente un coût de plusieurs centaines de francs pour une durée de quelques semaines. Par exemple, un cours d'allemand niveau A1 prodigué par l'Ecole-Club Migros de la Chaux-de-Fonds coûte CHF 655.50² pour 34.5 périodes, soit CHF 19.- par période. Si ce tarif était appliqué aux cours suivis par les élèves à l'école secondaire, cela représenterait un montant de CHF 15'960.- rien que pour l'allemand (840 périodes à raison de 40 semaines par an et 2 périodes par semaine de la 5^e à la 7^e, 3 périodes par semaine en 8^e, 4 périodes par semaine de la 9^e

¹ Selon (Mühlemann & Schöbi, 2016)

² Tarifs 2017, consultés le 4 janvier 2017 à l'adresse <http://www.ecole-club.ch/Offres/Langues/Cours-d-allemand>

à la 11^e, selon grille horaire Harmos³), ce que la grande majorité de la population ne pourrait se permettre.

Ces dernières années dans le canton de Neuchâtel, des efforts conséquents ont été fournis par les communes pour doter les salles de classe de vidéoprojecteurs et autres tableaux interactifs. Au final, de plus en plus de salles de classe disposent d'un tel équipement, qui facilite le rôle de l'enseignant en lui mettant à portée de main des ressources nouvelles. Il en résulte une modification du mode d'enseignement qui s'oriente vers une dimension numérique, offrant de nouvelles possibilités pour l'enseignement⁴.

Au vu de ce qui précède, il est donc à mes yeux primordial que l'enseignement langagier prodigué à l'école obligatoire soit qualitatif – et donc pérenne – afin que les montants investis par l'Etat dans la formation ne soient pas vains, et que les jeunes puissent se débrouiller dans les langues étudiées sans devoir avoir recours à des cours de langue d'instituts privés par la suite. De plus, la prolifération des nouveaux outils numériques en salle de classe dans les écoles neuchâteloises est une occasion rêvée de réfléchir à leur utilisation de manière à rendre l'enseignement prodigué encore plus efficace et pérenne.

Dans cette optique de mettre en place tout ce qui est possible pour que la matière vue en classe soit ancrée de manière durable dans la tête des apprenants – et ceci sur le long terme – j'ai parcouru divers ouvrages en lien avec mes didactiques d'enseignement – français, allemand et anglais – et me suis intéressé à un article présentant un dispositif qui favorise la rétention durable d'informations en salle de classe. Dans leurs travaux publiés en 2007, Butler & Roediger ont mis en évidence en milieu universitaire les avantages de tester des apprenants sur la matière traitée en fin d'une séquence, puis de répéter ces tests à diverses fréquences pendant les jours et semaines suivantes. Ce processus, connu sous le nom de *testing effect*, permettrait d'augmenter significativement la rétention d'informations à long terme. Dans une étude plus récente de 2011, Roediger III et al. ont obtenu des résultats similaires, cette fois au niveau d'une *middle school* de banlieue aux Etats-Unis, soit l'équivalent de notre école secondaire. La question de recherche était de déterminer si des tests à faible enjeu – ne comptant pas dans la note de l'apprenant ou dans une très faible proportion – pouvaient être utilisés pour améliorer d'une part la performance des apprenants lors de l'évaluation finale, d'autre part la rétention de l'information après un délai de plusieurs mois. Dans un cours de sciences sociales, ils ont intégré de manière non intrusive des quiz à des séquences didactiques d'enseignants

³ http://blogs.rpn.ch/jjrprofs/files/2010/09/k_CSR_HarmoS_DossierInformation.pdf, consulté le 4 janvier 2017

⁴ Selon Brochure « Compétences MITIC à l'école »

expérimentés. A chaque fois, certains contenus étaient testés et d'autres non. Ces quiz ont été administrés en amont de l'enseignement – à la suite d'une lecture préparatoire effectuée à la maison par les élèves –, après l'enseignement, et finalement 24 heures avant l'évaluation sommative validant le chapitre. Cette dernière comportait aussi bien des questions préalablement testées que des questions sur des sujets n'ayant pas été testés. Au final, les résultats de l'évaluation sommative ont révélé des scores supérieurs pour la matière préalablement testée par rapport à la matière non testée. Au niveau des types de tests proposés, ce sont les contenus testés sous forme de choix ouvert qui ont obtenu le meilleur score.

Par la suite, les étudiants ayant participé à l'étude ont également été testés lors de tests semestriels et annuels. Là aussi, la matière préalablement testée a été mieux remémorée, démontrant ainsi une rétention des informations sur le long terme.

Cette dernière étude avait pour cadre un cours de sciences sociales au cours duquel les élèves apprenaient à connaître différentes cultures du monde, ainsi que l'histoire de chacune d'elles. Bien que n'ayant pas été menée dans une classe de langue seconde, elle représente un excellent point de départ pour la recherche que je souhaite mener. En effet, les participants à l'étude étaient d'âge comparable à des élèves de 9^e Harmos dans le canton de Neuchâtel – 11 à 12 ans – et les tests ont été administrés de manière non intrusive en classe, sans modifier la pratique habituelle des enseignants, ce qui est également mon objectif. Enfin, les savoirs testés, au contraire d'autres études (Dirkx, Kester, & Kirschner, 2014; McDaniel et al., 2013; Mcdaniel, Agarwal, Huelser, Mcdermott, & Roediger, 2011) portant sur les mathématiques ou les sciences, appartiennent au domaine des faits et sont en ce sens plus proches à mes yeux de l'enseignement d'une langue seconde.

En plus de l'aspect lié à la rétention à long terme des informations, les tests de rappel sont un outil permettant l'amélioration des stratégies d'apprentissage des élèves. Selon Karpicke & Roediger (2008), les tests permettent de mesurer leur degré de compréhension de la matière, que ce soit en cours de séquence, ou pour la préparation d'une épreuve. Un bon résultat au test les confortera dans leur apprentissage, tandis qu'un résultat mitigé leur permettra de prendre conscience de leurs lacunes sur le sujet ; ils pourront donc travailler en conséquence ou poser les questions nécessaires en classe afin de les combler. Les tests sont également un moyen pour les élèves de discerner la matière importante des informations secondaires ; en étant testés régulièrement, les apprenants développent une « culture du test », leur permettant de déceler les contenus susceptibles d'être testés et peuvent se concentrer sur la mémorisation de ces contenus.

1.3 Etat de la question

Pour cette partie, je vais d'abord parler du fonctionnement de la mémoire, de la problématique de l'oubli et des facteurs qui favorisent la rétention d'informations en mémoire. Ensuite, je m'intéresserai aux études menées jusqu'à ce jour sur le *testing effect*, en mettant un accent particulier sur celles réalisées de manière non intrusive dans des classes réelles. Enfin, je traiterai de l'utilisation des quiz dans l'enseignement des langues secondes et des aspects liés à leur conception. Quant aux quiz informatisés, je parlerai brièvement de leurs avantages par rapport aux quiz papier et présenterai une sélection d'outils disponibles sur le marché.

1.3.1 Le fonctionnement de la mémoire

1.3.1.1 Types de mémoires

Selon Croisile (2009), la mémoire fait référence à trois processus chronologiques essentiels : apprendre des informations nouvelles, les conserver le plus longtemps possible, et les récupérer au moment opportun. On distingue deux types de mémoire : la mémoire à court terme et la mémoire à long terme.

Lors du processus d'apprentissage, les informations pertinentes sont dirigées vers la mémoire à court terme – également appelée mémoire de travail – dont les principales caractéristiques sont sa capacité limitée (7 plus ou moins 2 informations), sa volatilité (les informations y persistent 20 à 30 secondes) et sa sensibilité aux interférences (Croisile, 2009). Elle contient temporairement une courte série d'informations et permet de les comprendre, les répéter mentalement ou à voix haute, les manipuler (calcul, résolution de problèmes, vérification, décision), ou les analyser afin de les mémoriser définitivement. Cette dernière opération nécessite un encodage en profondeur des informations afin de les transférer en mémoire à long terme (Croisile, 2009).

Selon Croisile (2009), La mémoire à long terme est quant à elle constituée de quatre stocks :

- les mémoires perceptives des formes sensorielles : elles sont composées des traces des perceptions sensorielles indépendamment de leurs significations. Elles permettent de reconnaître une forme, une image ou un son avant de les identifier.
- la mémoire procédurale : elle permet l'apprentissage de procédures par la pratique répétée de tâches, comme des exercices sportifs ou le maniement d'outils et d'ustensiles. L'apprentissage initial s'effectue de manière consciente et lors d'un effort volontaire d'acquisition de la tâche, alors que la mise en œuvre ultérieure de l'habileté n'est plus consciente car le processus est inconsciemment automatisé. Les traces de ces apprentissages ne s'effacent pratiquement jamais : on peut citer comme exemple la pratique du vélo, ou le fait de savoir nager.
- la mémoire épisodique des souvenirs personnels : elle comporte les souvenirs personnels

vécus une seule fois, en un lieu déterminé et à un moment précis de la vie. Ces épisodes sont souvent associés à un contexte émotionnel et affectif particulier.

- la mémoire sémantique des faits culturels : elle est constituée de connaissances apprises plusieurs fois, qui n'ont plus d'indicateurs de temps et de lieu d'apprentissage et qui sont retenues indépendamment d'un contexte émotionnel et affectif. La mémoire sémantique est une mémoire collective composée de faits appris et partagés au sein de notre groupe social et culturel. Ces faits correspondent au langage (vocabulaire, grammaire) et aux savoirs sur le monde (histoire, géographie, célébrités, etc.).

C'est ce dernier stock mémoriel qui est mis à contribution dans le cadre de l'apprentissage d'une langue seconde, car les informations mémorisées résultent d'un apprentissage de faits appris à la suite de multiples répétitions (Thornbury, 2002).

1.3.1.2 Consolidation et rappel des informations

On parle de trace mnésique pour nommer l'empreinte laissée par une information mémorisée dans la mémoire à long terme. Il s'agit d'une modification du tissu nerveux constituant la base neurophysiologique de la conservation du passé dans la mémoire⁵.

Selon Croisile (2009), une information représentée par une trace mnésique solide sera plus aisément récupérée en mémoire qu'une trace mnésique estompée. La qualité de la trace mnésique est en premier lieu conditionnée par le codage initial. Celui-ci est favorisé par une attention soutenue et une comparaison en profondeur de l'information à enregistrer par rapport à ce qui est déjà su. Les émotions peuvent également faciliter la mémorisation en agissant comme des catalyseurs : des informations associées à un événement positif (par exemple une naissance ou un mariage) ou négatif (décès, accident) seront mieux retenues que des informations neutres. Ceci est en particulier valable lorsque ces émotions nous concernent directement : on se rappelle plus de son mariage que de celui d'un ami. Nos souvenirs, surtout s'ils sont positifs et agréables, seront mémorisés sur la durée. Ce phénomène est dû à une augmentation de l'activité de l'hippocampe qui se produit lorsque nous apprenons des informations à fort contenu émotionnel (Taconnat, 2012).

La mémorisation à long terme d'une information passe par la consolidation de la trace mnésique associée. Cette étape de consolidation d'une trace mnésique fraîche commence dans l'heure qui suit le premier apprentissage. Le sommeil contribue ensuite à la consolidation à long terme sous forme d'une

⁵ Définition de l'encyclopédie Universalis

réactivation des circuits neuronaux impliqués lors de l'apprentissage au cours de la journée précédente. Mais le véritable secret de la mémoire réside dans la répétition des apprentissages ; des informations récemment et régulièrement utilisées seront mieux consolidées et donc plus rapidement retrouvées : ainsi, « se rappeler fréquemment un savoir ou un souvenir revient à les réapprendre » (Croisile, 2009).

Au moment de récupérer une information en mémoire, notre cerveau réagit différemment selon la qualité de l'encodage initial (Croisile, 2009). En effet, « plus une information aura été élaborée, organisée, structurée, plus elle sera facile à retrouver » (Croisile, 2009). En ce sens, une étude de 1999 (Bower, Clark, Lesgold, & Winzenz) a démontré que l'organisation en catégories hiérarchisées facilite grandement le rappel d'un nombre élevé d'informations. Au cours d'une expérience de mémorisation de 112 mots, certains sujets les ont appris mélangés au hasard, tandis que d'autres les ont appris sous une forme logique hiérarchisée (par catégories). Après 3 essais, 100% des mots ont été rappelés par le groupe « organisé », contre seulement 47% pour le groupe « mélangé ».

Toujours selon Croisile (2009), Le rappel d'une information en mémoire peut prendre trois formes : rappel libre, rappel indicé, et reconnaissance parmi plusieurs solutions proposées. Lors de la phase de mémorisation, on incite le sujet testé à faire un lien entre chaque mot et sa catégorie (par exemple : « girafe est un animal »). Dans le cas d'un rappel libre, le sujet cherche à se rappeler le terme en tentant de se remémorer quels étaient les mots. On parle de rappel indicé lorsqu'un indice lié à la catégorie est fourni : « il y avait un nom d'animal ». Enfin, la reconnaissance implique de reconnaître le mot mélangé à des éléments distracteurs, par exemple dans une liste : lion / buffle / girafe. Au final, « l'efficacité du rappel est, par ordre décroissant, le rappel libre, le rappel indicé, et enfin la reconnaissance » (Croisile, 2009). Le cas du rappel libre sera discuté plus en détails lors du chapitre traitant du *testing effect*.

1.3.1.3 Le mécanisme de l'oubli

Même avec la meilleure volonté du monde, toute information finit par être oubliée si elle n'est pas réactivée en mémoire (Thornbury, 2002). L'oubli est rapide au départ, puis diminue progressivement, selon la courbe conceptualisée par Ebbinghaus et présentée plus haut. On estime que pendant les 24 heures qui suivent l'apprentissage, environ 80% de l'information est perdue. Ainsi, la mémoire à long terme peut être vue non pas comme un espace de stockage homogène comme on pourrait le penser, mais plutôt comme un continuum allant du « vite oublié » au « jamais oublié » (Thornbury, 2002), le grand défi pour les apprenants étant de transformer la matière de la première catégorie à la seconde.

Pour palier cet oubli massif, il a été démontré qu'il est possible de réduire drastiquement la quantité d'informations oubliées via des rappels périodiques (Nicolas, 1992). Dans le cadre de l'apprentissage du vocabulaire en langues secondes, Thornbury (2002) a démontré que les apprenants ont besoin de multiples expositions aux mots pour les mémoriser et que les mots qui sont appris sur plusieurs sessions espacées sont mieux retenus. De plus, pour être retenus, les mots doivent être extraits régulièrement de la mémoire de l'apprenant.

1.3.1.4 Stratégies de mémorisation à long terme

Dans le cadre de l'apprentissage des langues secondes, l'évolution des connaissances en matière de stockage et d'extraction d'informations en mémoire a directement influencé les méthodes d'enseignement ; de relativement marginal par le passé où les méthodes donnaient la priorité à l'acquisition des structures grammaticales, l'apprentissage du vocabulaire est devenu un objectif à part entière dans les années 70, avec la progression de l'approche communicative (Thornbury, 2002). Suite à la découverte que le cerveau ne stocke pas les mots au hasard ou sous la forme d'une liste, mais d'une manière organisée et interconnectée, les listes de vocabulaire ont progressivement évolué, passant de listes alphabétiques à des listes regroupant des mots en contexte, centrés autour d'un même sujet (Thornbury, 2002). Afin de faciliter la rétention des éléments à long terme, Thornbury (2002) propose d'agir sur les axes ci-après, sur la base des découvertes de la recherche :

- répétition : un mot revu plusieurs fois, dans des contextes différents, est mieux mémorisé.
- récupération : le processus d'extraction d'éléments en mémoire contribue à leur mémorisation.
- espacement dans le temps : permet de lutter contre la courbe de l'oubli.
- utilisation : le fait d'utiliser un terme dans son vocabulaire productif augmente sa pérennité en mémoire.
- image mentale : associer un mot à une ou plusieurs images contribue à augmenter le nombre de connexions mentales pour ce mot.

Ces divers éléments, s'ils sont pris en compte par l'enseignant et intégrés dans son enseignement, favorisent la mémorisation à long terme des contenus langagiers.

1.3.2 Le testing effect

Dans la plupart des environnements éducatifs, les tests sont employés comme moyen d'évaluation dans le but de décerner des notes. Cet usage éclipse cependant souvent une autre fonction des tests

qui est particulièrement pertinente pour l'éducation, à savoir la promotion de l'apprentissage. En effet, de nombreuses études dans le domaine de la psychologie cognitive ont démontré l'impact favorable des tests sur la rétention de la matière testée ; ce phénomène est connu sous le nom de *testing effect* (Roediger & Karpicke, 2006).

Les études sur le *testing effect* consistent à étudier deux groupes étudiant la même matière selon la méthode expérimentale ; un groupe passe un ou plusieurs tests après exposition à la matière, alors que le second groupe – le groupe témoin – n'a soit plus affaire à la matière, soit relit la matière le même nombre de fois que le groupe test est testé. Tous les participants sont ensuite testés lors d'un test final de rétention quelque temps après.

1.3.2.1 Un sujet d'étude centenaire

L'idée d'utiliser les tests comme outil d'apprentissage en salle de classe remonte au début du XX^e siècle avec notamment les travaux de Gates en 1917, mais Jones fut le premier chercheur à transférer les découvertes de laboratoire dans une salle de classe en 1923-1924. Alarmé par la faible rétention du contenu de ses cours, il fit passer un test court composé de lacunes à compléter et de réponses courtes immédiatement après un cours d'une heure, puis effectua des tests similaires après des délais variant de 3 jours à 8 semaines, afin de mesurer la quantité d'informations oubliées. Les résultats montrèrent que la quantité d'informations retenues après 8 semaines dans le cas d'un test préalable était plus grande qu'après seulement 3 jours si aucun test n'avait été effectué en amont. Jones en conclut que les tests étaient une méthode efficace pour augmenter la rétention de contenu de cours et précisa que ceux-ci devraient être effectués juste après afin d'en maximiser les effets (Spitzer, 1939).

Par la suite, de nombreuses études investiguèrent les avantages du *testing effect*. Elles suggéraient que les tests en milieu scolaire influençaient positivement la rétention à long terme de contenu de cours, et cela dans de nombreuses disciplines, notamment la physique ou l'apprentissage des langues (Semb & Ellis, 1994). Apparemment, les effets de ces tests augmentent avec le nombre de tests passés et lorsque ceux-ci comportent du feedback, sous la forme d'indications en réponse à des réponses erronées (Roediger & Karpicke, 2006b). Concernant le feedback, de meilleurs résultats sont obtenus lorsqu'il est différé par rapport à un feedback immédiat (Roediger & Karpicke, 2006b). Ces études furent critiquées car aucune ne manipulait les tests comme une variable indépendante tout en contrôlant les variables confondantes (méthode expérimentale) ; en effet, le fait de conduire les études dans de vraies classes entraînait de nombreux facteurs incontrôlés, par exemple la possibilité d'étudier la matière en-dehors du cadre scolaire. Quelques études expérimentales en laboratoire manipulèrent les tests comme

variable indépendante, cependant avec des contenus très simples, tels que des listes de mots (Butler & Roediger, 2007). Ce n'est qu'en 2007 que McDaniel et al. conduisirent une étude similaire sur du contenu complexe, à savoir un cours universitaire sur le fonctionnement du cerveau. Il en résulta que la rétention était supérieure en effectuant un test de réponses courtes juste après le cours, par rapport aux groupes ayant eu un questionnaire à choix multiples ou des phrases clés à lire. Un problème de cette étude était cependant que les étudiants étaient jugés seulement sur des faits périphériques en lieu et place de la matière centrale qui était testée dans le cadre du cours, ceci en raison de craintes relatives au potentiel impact des tests sur les notes des étudiants (Roediger III et al., 2011).

En 2007 également, Butler & Roediger menèrent une étude sur les avantages d'effectuer des tests dans un environnement scolaire simulé. L'idée était de combiner le contrôle expérimental d'un laboratoire avec la matière d'un cours universitaire, en l'occurrence un cours d'histoire de l'art. Dans un souci de réalisme, un intervalle de rétention relativement long (1 mois) a été utilisé, ce délai représentant plus ou moins le temps pendant lequel des étudiants sont appelés à se rappeler des informations avant une épreuve. Deux groupes furent examinés en parallèle : un groupe contrôle et un groupe test. L'objectif de l'étude était d'investiguer dans quelle mesure différents types de tests (questionnaires à choix multiples et questions courtes) et la communication d'un feedback (réponses correctes communiquées ou non) augmenteraient la rétention de la matière des cours.

Les résultats de l'étude furent convaincants. Butler & Roediger démontrèrent que le fait d'effectuer une activité d'apprentissage, peu importe laquelle des 3 proposées, résultait en une meilleure rétention de la matière. Le test le plus bénéfique fut le test de réponses courtes, qui produit une meilleure rétention de la matière que l'étude d'un résumé du cours ou le passage d'un questionnaire à choix multiples. Néanmoins, l'inconvénient de cette étude résultait dans le fait qu'elle se déroulait en laboratoire et non dans une véritable salle de classe.

1.3.2.2 Le testing effect dans un environnement scolaire réel

Il a fallu attendre 2011 (Roediger III et al.) pour qu'une étude cherche à prouver l'efficacité des tests à différents niveaux du système éducatif. Ils menèrent sur une durée de 3 semestres des expériences dans lesquelles ils utilisèrent le véritable contenu du cours, ainsi que les tests sommatifs et examens semestriels du cours comme mesures dépendantes. Ils utilisèrent les textes traités en classe ainsi que l'ensemble des documents précédemment créés par l'enseignant, dans le but de maintenir les habitudes de la classe. Pour la réalisation des quiz, des questionnaires à choix multiples développés par l'enseignant ont été utilisés, alors que les tests finaux comportaient des questions à choix multiples et des réponses courtes. Les questions des quiz couvraient la même matière que pour le test final.

Une première expérience a été menée sur 142 élèves de 6^e année d'une *middle school* aux Etats-Unis, dans le cadre d'un cours de sciences sociales. L'objectif de cette expérience était de voir si un *testing effect* était mesurable entre des éléments préalablement testés et d'autres non testés. La matière utilisée était celle de quatre chapitres du livre de sciences sociales et traitait des différentes cultures autour du monde, et de leur histoire (Egypte ancienne, Mésopotamie, Inde et Chine). En amont de chaque leçon, un quiz initial sur la matière testée plus tard a été administré aux élèves, puis un second directement après la leçon. Approximativement 2 jours plus tard, les élèves passèrent un test de révision sur les éléments testés. A chaque fois, un assistant de recherche était présent en classe pour administrer les quiz oralement et visuellement à l'aide d'un système de boîtiers électroniques. Lors de chaque test, la moitié des objectifs de chaque chapitre furent testés à l'aide de questionnaires à choix multiples, alors que l'autre moitié n'était pas testée. Après avoir répondu à chaque question, les élèves reçurent un feedback immédiat leur indiquant la bonne réponse.

Le test de mesure de la rétention fut administré par l'enseignant en général deux jours après le test de révision, au format papier et crayon. Celui-ci contenait un examen de type QCM comportant tous les éléments testés et ceux non testés. Les questions étaient les mêmes que celles des quiz précédents, présentés dans le même ordre que la matière vue en cours.

Les résultats de cette étude ont révélé que la performance des élèves pour les éléments préalablement testés était plus grande (94%) que pour les éléments non testés (81%), démontrant l'existence d'un *testing effect*. Par la suite, la rétention à long terme a également été mesurée lors d'examens de fin de semestre et de fin d'année, également sous la forme de QCM. Là aussi, la performance pour les éléments testés (79%) a été supérieure aux éléments non testés (67%).

Afin de vérifier que ces résultats n'étaient pas simplement dus à une étude supplémentaire de la matière (en raison du feedback fourni lors des tests), une seconde expérience a été menée. L'idée était de voir si la répétition des tests permettait un gain par rapport à un nombre de révisions de la matière identiques. Les mêmes participants ont été testés selon le même protocole, cette fois sur 3 chapitres additionnels du livre de sciences sociales (Rome ancienne, le Moyen-Âge, l'Afrique). Les variations suivantes ont cependant été apportées pour chacun des quiz (pré-test, post-test et test de révision) : un tiers des objectifs de chaque chapitre ont été testés au format QCM, un tiers a été présenté pour lecture, et le dernier tiers n'a pas été testé. A nouveau, les résultats ont démontré l'existence d'un *testing effect* pour la performance à l'examen de fin de chapitre pour les items testés (91%) par rapport à ceux lus (83%) et non testés (81%). Ces conclusions, en plus de confirmer que les tests augmentent

la rétention par rapport à la révision, ont démontré que la relecture de la matière n'augmente pas significativement la performance, ce que d'autres études avaient également souligné (Karpicke & Roediger, 2008 ; Brown, 2014).

Une troisième expérience menée dans le cadre de cette étude avait pour but de voir si des résultats similaires pouvaient être observés en remplaçant les questions à choix multiple par des réponses courtes. Celle-ci confirma que la performance des éléments testés au format réponses courtes était significativement meilleure (84%) que celle des éléments non testés (75%). Des conclusions similaires avaient déjà été obtenues dans une précédente étude (McDaniel, Anderson, Derbish, & Morrisette, 2007), qui a montré que le fait d'élaborer soi-même les réponses et d'être forcé de composer des phrases avec ses propres mots permettait une meilleure rétention de l'information.

Au final, les 3 expériences menées ont permis de révéler les effets positifs du rappel d'informations en mémoire à l'aide de quiz dans le cadre scolaire, sur du matériel issu du programme. De plus, des questionnaires transmis aux élèves testés en fin d'année scolaire ont révélé que 97% d'entre eux considéraient l'utilisation des boîtiers de test comme augmentant leur apprentissage, et 65% ont rapporté que l'utilisation des boîtiers diminuait leur anxiété par rapport aux tests. En outre, les élèves ont perçu l'utilisation des boîtiers de test en classe comme bénéfique et appréciable.

Bien que dans cette étude des questions similaires aient été utilisées pour le test de révision et le test final, d'autres études ont démontré que les tests comportant des questions modifiées augmentent également la rétention (McDaniel, Thomas, Agarwal, Roediger, & McDermott, 2011), ainsi qu'un transfert de connaissance vers de nouvelles situations (Butler, 2010).

1.3.2.3 Importance du feedback

Un aspect particulier des études sur le *testing effect* concerne la notion de feedback. Par feedback, on entend le retour donné à un apprenant à la suite d'une action ; cette information lui permet de savoir si la réponse donnée est correcte ou non. Après avoir été considéré comme une variable pour certaines recherches (notamment Butler & Roediger, 2007), il a pu être démontré qu'il jouait un rôle important dans la mémorisation des informations (R. C. Anderson, Kulhavy, & Andre, 1971; Butler, Karpicke, & Roediger, 2007, 2008; Butler & Roediger, 2008; Lhyle & Kulhavy, 1987). Dans les études les plus récentes, dont celle de Roediger et al. (2011), le feedback fait partie intégrante du processus de test.

Concernant la forme à donner à ce feedback, Bravard (2005) distingue deux composantes :

- la composante dite de vérification qui fournit un jugement d'exactitude/erreur permettant à l'apprenant de vérifier la pertinence de son action.
- la composante explicative qui apporte d'autres informations complémentaires. Les erreurs peuvent ainsi être comprises et corrigées par l'apprenant.

Le feedback doit être court et contenir juste assez d'informations pour permettre à l'apprenant de comprendre l'erreur commise. Si toutefois la matière à apprendre est plus complexe une rétroaction plus étendue peut être justifiée (Bravard, 2005).

De nombreuses études récentes questionnent cependant l'aspect du délai entre la réponse et le feedback (Metcalf, Kornell, & Finn, 2009; Stefanou & Révész, 2015; Li, Zhu, & Ellis, 2016). Dans leur étude de 2009, Metcalfe, Kornell & Finn ont démontré qu'un feedback différé donnait de meilleurs résultats pour des élèves américains de 6^e année (âgés de 11-12 ans) pour une tâche de mémorisation de paires mot-définition d'un cours de sciences sociales, par rapport à un feedback immédiat :

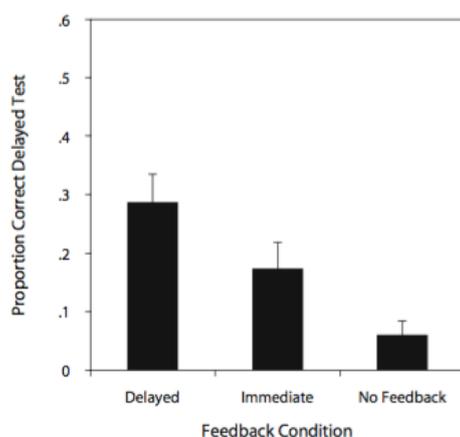


Figure 3: source: Metcalfe, Kornell & Finn (2010), p. 1082

Dans cette étude, le feedback aux questions posées lors d'une séance survenait la séance suivante, soit dans un délai variant entre 1 et 4 jours. Les auteurs expliquent ces résultats par le fait qu'un feedback différé donne l'opportunité à l'élève de se remémorer la matière après un certain délai, ce qui en soi représente un rappel.

Lors d'une autre expérience sur une population adulte composée d'étudiants universitaires américains, la même étude a révélé des résultats identiques entre les deux conditions de feedback.

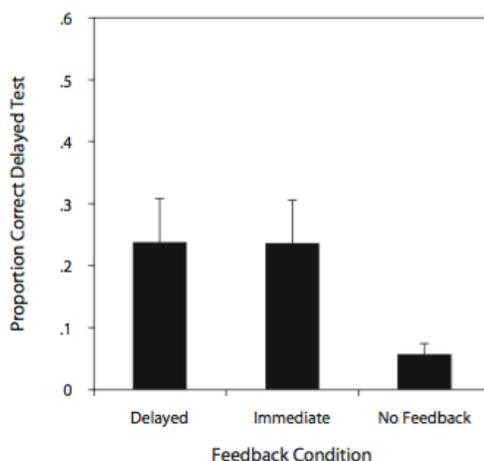


Figure 4: source: Metcalfe, Kornell & Finn (2010), p. 1084

Ces résultats sont néanmoins à considérer avec précaution, car l'expérience a été menée en laboratoire plutôt que dans une véritable classe. La matière testée était aussi différente : il s'agissait de paires mot-définition d'un cours de niveau universitaire.

Dans une autre étude (Brown, Roediger et Mc Daniel, 2014), la durée entre la réponse au test et le feedback a été ramenée à quelques minutes seulement ; dans ce cas également, le résultat a été meilleur qu'avec un feedback immédiat. L'hypothèse avancée par les chercheurs pour expliquer ce résultat est que si le feedback n'est pas fourni directement, l'effort demandé pour la remémoration de l'information est plus important, ce qui aurait pour effet de consolider la mémorisation. Ce résultat est particulièrement intéressant dans le cadre de l'enseignement, car l'enseignant peut aisément différer son feedback de quelques minutes au cours de sa leçon, ou profiter des dernières minutes pour donner ce feedback, pour conclure la leçon.

Dans le contexte plus spécifique de l'apprentissage des langues secondes, qui m'intéresse dans le cadre de mon travail, Li, Zhu & Ellis (2016) ont conclu qu'un feedback immédiat donne des résultats légèrement supérieurs par rapport à un feedback différé dans le cadre de l'acquisition de nouvelles structures grammaticales. Leur étude s'intéressait à l'acquisition de la structure du *past passive* en anglais, dans le cadre d'une école de langue chinoise (participants âgés entre 13 et 15 ans). Lorsque les deux types de feedback ont été comparés, le feedback immédiat a donné de meilleurs résultats 8 fois sur 10. De plus, les effets positifs du feedback immédiat se maintenaient après 2 semaines, alors que les effets positifs du feedback différé (par rapport au groupe témoin) ne perduraient pas après le post-test. Leur étude souligne cependant que leurs conclusions ne permettent pas de donner un

support sans équivoque au feedback immédiat et que des recherches complémentaires sur la question sont nécessaires.

1.3.2.4 Conclusions sur le testing effect

Au vu de ce qui précède, les tests de rappel effectués sous la forme de quiz en salle de classe se profilent comme des candidats sérieux pour lutter contre le phénomène de l'oubli. En effet, la recherche démontre que quand le cerveau doit travailler, la mémorisation est meilleure ; plus l'effort de rappel d'information en mémoire est important, plus l'apprentissage est renforcé par le rappel (Brown, 2014).

Comme le soulignent McDaniel et al. (2011), les quiz s'intègrent aisément dans une pratique enseignante quotidienne et se basent sur la matière et les moyens d'enseignement existants. De plus, ils n'impliquent que des changements mineurs pour l'enseignant, qui doit les préparer et réserver un peu de temps pour leur passation en classe.

Au niveau du feedback, il semble que la question du moment auquel le donner soit un thème âprement discuté et sujet à controverse. Dans mon cas, les travaux de Metcalfe, Kornell & Finn (2009) sur une population d'adolescents me pousseraient à opter pour un feedback différé, alors que ceux de Li, Zhu & Ellis (2016) favoriseraient le feedback immédiat pour les langues secondes. Pour mon étude, j'ai choisi de faire un compromis et fournir un feedback différé de quelques minutes par rapport aux réponses aux questions ; un retour sera donné à l'apprenant à la fin du quiz, au moment de la communication de ses résultats.

1.3.3 Les quiz dans l'apprentissage des langues

Cette section a pour objectif de présenter la manière dont les quiz sont utilisés dans l'apprentissage des langues secondes, et les avantages qu'ils apportent aux élèves et enseignants, en plus d'améliorer la rétention. Les fondements de base liés à leur conception et à la définition des items seront passés en revue, tout comme deux types de question qui nous intéressent particulièrement dans cette étude : les questions de type QCM et les réponses courtes. Enfin, l'apport de l'informatique pour la passation des quiz et ses incidences seront présentés, tout comme une sélection d'outils présents sur le marché.

1.3.3.1 Présentation

Un quiz représente un jeu sous la forme d'un concours par questions et réponses, au terme duquel un résultat est communiqué⁶. Dans l'apprentissage des langues secondes, les quiz sont largement répandus et permettent notamment de tester des compétences telles que la compréhension orale et écrite.

Dans le domaine des langues, les quiz sont souvent composés de types de questions différents, en fonction de l'objectif langagier recherché. Parmi les plus courants, on dénombre les types suivants :

- Vrai/faux : l'étudiant doit répondre à une question par « vrai » ou « faux ».
- Questions à choix multiples (QCM) : l'étudiant doit répondre à une question à partir d'un choix multiple de réponses. On dénote deux types de QCM, ceux où une seule bonne réponse est permise, et ceux où de multiples réponses sont permises.
- Appariement : une liste de propositions est affichée à côté d'une liste de réponses. L'étudiant doit appairer chaque proposition à la réponse correspondante.
- Composition (ou essai) : l'étudiant est appelé à répondre à une question sous forme de composition. Selon les cas, le nombre de caractères ou de mots requis est indiqué.
- Lacune : ce type de question permet de proposer des phrases ou textes à trous dans lesquels s'insèrent des réponses. Ces questions peuvent être de types liste à choix multiple ou réponse courte.
- Réponse courte : l'étudiant doit répondre à une question en quelques mots. Il peut y avoir plusieurs bonnes réponses possibles.

Dans le cadre des travaux de Roediger III et al. (2011), les deux types de questions qui ont été utilisés dans les quiz ont été le QCM et les réponses courtes. Nous allons tenter de répliquer en partie son dispositif et donc nous cantonner par la suite à ces deux types de questions.

1.3.3.2 Avantages des quiz

En plus des avantages liés à la meilleure rétention des informations, Roediger, & Smith (2011) identifient 9 bénéfices des quiz dans un environnement éducatif :

1. Ils permettent d'identifier des lacunes chez les élèves.
2. Grâce à eux, les apprenants apprennent plus lors de la phase d'apprentissage suivante.
3. Ils permettent une meilleure organisation des connaissances.

⁶ Définition du Larousse.

4. Ils améliorent le transfert de connaissances vers d'autres contextes.
5. Ils peuvent faciliter la récupération d'informations non testées.
6. Les quiz ont un impact positif sur les capacités métacognitives des apprenants quant à la révision ; Roediger III et al. (2011) soulignent également ce point en concluant que les élèves peuvent ainsi prendre conscience de ce qu'ils savent et ce qu'ils ne savent pas, afin de concentrer leurs efforts futurs sur l'information qu'ils ne savent pas. De plus, Agarwal et al. (2008) indiquent que la pratique régulière de tests permet d'augmenter le jugement ou les prédictions des élèves dans leurs apprentissages ; suite aux tests, les élèves disposent d'un jugement métacognitif plus précis concernant la matière potentiellement testée.
7. Des quiz fréquents encouragent les élèves à travailler.
8. Ils peuvent également être utilisés par les élèves pour s'auto-évaluer.
9. Enfin, ils offrent un feedback aux enseignants sur le niveau de compréhension des élèves.

Ce dernier point va dans le sens des travaux d'Hattie (2008). Dans un ouvrage couvrant plus de 15 années de recherches sur les influences en matière de réussite scolaire, il émet le postulat que pour que chaque étudiant puisse progresser dans son cursus scolaire, les enseignants doivent être au courant de ce que chaque élève sait et pense pour pouvoir construire des enseignements pertinents et sensés, et fournir un feedback approprié à chacun (Hattie, 2008). Par la suite, dans un autre ouvrage destiné aux enseignants, il souligne qu'en plus de rendre l'apprentissage des élèves visible pour les enseignants, il est également nécessaire de rendre l'enseignement visible pour les élèves, de manière à ce qu'ils puissent développer une envie d'apprendre et la capacité d'être leur propre enseignant, dans l'optique d'un apprentissage à vie (Hattie, 2011).

Butler et Roediger (2007) concluent quant à eux que des tests fréquents peuvent augmenter la régularité des révisions des étudiants et aider à diminuer leur anxiété en situation de test, ce qui permettrait d'améliorer leur performance lors d'évaluations sommatives. Les auteurs encouragent les enseignants à incorporer les tests dans leur routine quotidienne, le temps investi pour effectuer un quiz étant moindre en comparaison avec les avantages d'une telle pratique.

Ces dernières années, l'essor de l'informatique et la démocratisation de l'accès à internet ont engendré une explosion du nombre de quiz disponibles. Cet aspect est le sujet du chapitre suivant.

1.3.3.3 L'apport de l'informatique

Dans l'apprentissage des langues secondes, les quiz informatisés sont très répandus ; on ne dénombre plus le nombre d'éditeurs de méthodes qui en proposent en ligne. Selon Bravard (2005), une de raisons principales de cet essor consiste en l'introduction des TIC, qui ont permis un traitement immédiat de l'erreur et de la rétroaction immédiate que l'apprenant peut recevoir : le feedback. Ainsi, l'apprenant n'a plus à attendre la phase de correction pour pouvoir passer à la phase de remédiation. De plus, un des atouts largement reconnus de l'utilisation des TIC consiste en l'individualisation des rythmes des apprenants ; il dispose de tout le temps nécessaire pour réfléchir et peut ainsi prendre confiance en lui, en se libérant d'une certaine pression. Seul derrière son écran, l'apprenant peut oser, sans crainte de voir sa performance comparée à celle des autres.

Harmer (2015) souligne également que beaucoup de tests publics sont maintenant administrés de manière digitale. De cette manière, l'étudiant a la possibilité de recevoir une note automatiquement, ce qui signifie que la correction est effectuée numériquement et est en ce sens plus fiable qu'une correction humaine. Les tests numériques informatisés offrent également une fiabilité absolue en matière de conditions de test, car chaque candidat va recevoir un traitement identique. Cet aspect lié à la correction automatisée est particulièrement intéressant pour l'enseignant, qui n'a plus à corriger manuellement les quiz de ses élèves. Il peut ainsi se concentrer sur la création de nouveaux quiz ou la mise à disposition de quiz existants sur un sujet donné.

Un autre aspect positif de la numérisation des quiz concerne leur démocratisation. Ces dernières années, avec l'émergence d'outils facilitant leur création et de plateformes dédiées à l'enseignement – notamment Quia.com ou Learningapps.com –, de nombreux quiz réalisés par des enseignants pour leurs classes, puis mis à disposition de la communauté, ont vu le jour (Bravard, 2005). Dans le canton de Neuchâtel, la plateforme iClasse, développée par le service cantonal de l'enseignement, permet d'utiliser des ressources numériques, de les organiser en parcours et de les distribuer aux élèves. La création de quiz n'est ainsi plus réservée aux éditeurs de méthodes de langues ou écoles de langues, mais accessible à tout un chacun.

A mes yeux, les quiz informatisés possèdent encore les avantages suivants par rapport à leurs équivalents – et précurseurs – au format papier :

- l'accessibilité : le fait que les quiz soient hébergés sur des serveurs web les rend accessible en-dehors de la salle de classe, à tout moment de la journée, p.ex. sur les smartphones, dans le train, en toute discrétion.

- les essais différés : avec les quiz numériques, il est possible de repasser le test en tout temps, afin de mesurer la rétention des informations, vérifier un progrès, ou réactiver ses connaissances.

Bien que de nombreuses études sur le testing effect aient utilisé des tests au format papier, les études les plus récentes ont eu recours au format numérique. En ce sens, les travaux de Clariana & Wallace (2002) ont montré que pour deux groupes ayant effectué les mêmes tests sur des supports différents (numérique et papier), le groupe les ayant effectués sur ordinateur a obtenu de meilleurs résultats que l'autre. Les auteurs mentionnent ainsi un possible *test mode effect* (effet lié au type de support utilisé pour le test). Les travaux plus récents de Chua & Don, (2013) confortent l'existence de ce *test mode effect* et révèlent en outre que les tests réalisés sur ordinateur augmentent la motivation intrinsèque et la self-efficacité des apprenants, tout en diminuant leur anxiété.

Au vu de ce qui précède, il me semble intéressant d'utiliser des quiz informatisés. Dans de nombreuses écoles romandes, des réflexions sont menées à l'interne afin d'autoriser l'utilisation des smartphones personnels des élèves à des fins pédagogiques, j'ai donc décidé de surfer sur cette vague et d'utiliser ces supports numériques pour ma recherche. De plus, l'utilisation de quiz informatisés répond au critère énoncé au chapitre 1.1 quant au souhait de mettre en place un dispositif ne nécessitant pas de support papier.

Parmi la multitude d'outils disponibles sur le marché pour la création de quiz informatisés, j'ai décidé d'effectuer une sélection répondant aux critères cités plus hauts. Cette sélection est détaillée ci après.

1.3.3.3.1 Sélection d'outils disponibles sur le marché

De nombreux outils existent sur le marché pour la réalisation de quiz informatisés. Parmi eux, certains nécessitent d'être installés de manière « standalone » sur un ordinateur, alors que d'autres sont des applications orientées web. Dans l'optique de pouvoir accéder au dispositif via n'importe quel périphérique connecté et d'éviter toute modification ou installation sur les ordinateurs, j'ai décidé de privilégier la seconde catégorie. Ce choix implique la nécessité d'une connexion à Internet pour accéder aux quiz, mais cela m'est apparu comme un maigre inconvénient comparé aux avantages d'une telle solution.

Un autre critère essentiel à mes yeux concernait l'interface de l'outil ; celle-ci devait être « responsive », c'est-à-dire être utilisable aussi bien sur un ordinateur de bureau, que sur une tablette ou un smartphone, à l'image de la majorité des sites Internet ou des applications orientées web.

Les autres critères que j'ai décidé de prendre en compte étaient les suivants :

- Inclusion de médias : possibilité d'intégrer des images, du son ou des vidéos aux quiz.
- Rétroaction : moyens et délais pour donner un feedback. J'ai choisi ici des outils permettant de donner un feedback immédiat ou légèrement différé (feedback communiqué en fin de quiz, en même temps que les résultats), pour les réponses erronées comme correctes.
- Notation : l'outil doit permettre de communiquer à l'élève un résultat global quant au quiz passé, sous forme de pourcentage de bonnes réponses, d'appréciation, ou de note.
- Authentification : nécessité ou non de se créer un compte pour passer le quiz. Pour des raisons de protection des données et de facilité de mise sur pied, j'ai opté pour des outils ne nécessitant pas d'authentification.
- Restriction d'accès : quiz privé, accessibles via un lien, ou publics, librement accessibles par un moteur de recherche. J'ai décidé d'opter pour des outils permettant de rendre les quiz privés.
- Possibilité d'exporter facilement les résultats des apprenants et d'obtenir des statistiques suite à la passation du test (taux de réussite moyen, taux de réussite par question, temps moyen pour la passation, etc).
- Suivi des connexions et activités des apprenants : certains outils offrent des rapports détaillés des activités de chaque apprenant lorsqu'ils sont connectés via leur compte utilisateur, permettant ainsi par exemple de connaître les élèves qui utilisent activement le dispositif. Dans mon cas, comme les outils ne nécessitent pas d'authentification, je procède à l'identification de l'utilisateur en lui demandant son prénom en début de quiz, ce qui est une information suffisante pour mon étude. Dans un cas de litige (un élève ayant par exemple utilisé le prénom d'un autre pour compléter un quiz), l'adresse IP du périphérique utilisé permet de démasquer l'auteur.
- La maintenance de la plateforme doit être externalisée et opérée par une fondation ou une entreprise.
- Possibilité de personnaliser le quiz par l'ajout d'un texte introductif ou de conclusion, de modifier la couleur ou la typographie de l'interface.
- Possibilité de changer l'ordre des questions après la création, ou d'afficher les questions dans un ordre aléatoire.
- Possibilité d'afficher les réponses aux questions de type QCM de manière aléatoire.
- Affichage de l'avancement du quiz, soit en pourcentage, soit en indiquant le nombre de questions restant à répondre.
- Réutilisation possible d'un quiz existant ou de questions d'un quiz.

Enfin, les critères énoncés au chapitre 0 (dispositif peu intrusif et pouvant s'intégrer à un enseignement existant, nul besoin de connaissances techniques poussées ou d'équipe technique pour la mise en place du dispositif, dispositif peu chronophage pour l'enseignant, coût moindre et pas d'achat de matériel spécifique nécessaire) ont également été prépondérants dans ma sélection.

Au final, j'ai décidé de sélectionner 2 outils, qui répondaient aux différents critères. Ils sont présentés ci-après :

[Google Forms \(forms.google.com\)](https://forms.google.com)

Bien que conçu à la base pour la création de formulaires en ligne, l'outil de Google permet également de réaliser des quiz en cochant dans les paramètres du formulaire l'option « Make this a quiz ». Son utilisation nécessite un compte compatible Google.

Avantages

L'un des atouts majeurs de cet outil est sa gratuité et son potentiel d'évolution. A l'instar de nombreux de ses services, Google a prévu la possibilité d'ajouter des extensions à Google Forms, qui peuvent être développées par tout un chacun — pour autant qu'on ait quelques notions de programmations — selon ses besoins.

Google Forms dispose également d'une grande variété de types de questions. Pour les questions de type QCM d'abord, il est possible d'utiliser des choix multiples, des cases à cocher, ou encore des listes déroulantes. Parmi les questions demandant à l'utilisateur d'entrer du texte, il y a le choix entre les réponses courtes et les paragraphes, qui permettent l'écriture d'essais. Quatre autres types de questions sont encore disponibles, à savoir l'échelle linéaire, la grille de choix multiples, la date et l'heure.

Concernant les médias supportés, là encore l'impact de Google se fait sentir. Il est en effet possible d'ajouter du texte, des images ou des vidéos à n'importe quel endroit du quiz. L'ajout d'images peut s'effectuer par téléchargement d'une image personnelle, par l'utilisation de l'appareil photo, en spécifiant son URL, par Google Drive ou encore par une recherche Google. Pour les vidéos, il est possible de spécifier l'URL ou d'effectuer une recherche sous Youtube.

Au niveau de l'interface, il est possible de modifier la couleur du fond parmi une palette de couleurs, de choisir un thème parmi de nombreuses catégories, ou d'uploader une image personnelle. Des sections peuvent être ajoutées afin de provoquer le passage à une nouvelle page et une barre de progression indique l'état d'avancement du quiz.

L'ordre des questions peut être défini comme aléatoire et le choix de la notation du quiz peut être immédiat ou après une analyse manuelle (par exemple pour des questions de type réponses courtes ou paragraphe). Pour les questions de type QCM, il est possible de différencier le feedback selon si la réponse est correcte ou erronée.

Concernant les résultats, l'interface est très conviviale et permet de visualiser les résultats globaux ou détaillés en un clin d'œil. Différents graphiques permettent pour chaque question de se faire une idée des performances des apprenants. Les résultats peuvent également être exportés au format CSV ou directement dans une feuille de calcul Google.

Inconvénients

Pour conserver une information sur le preneur du quiz afin de pouvoir l'identifier, il est nécessaire d'ajouter un champ texte pour l'identité. La récolte de cette information n'est pas prévue de base où seule la récolte d'une adresse e-mail est possible.

Il n'est non plus possible d'intégrer uniquement du son à un quiz, ou d'utiliser une vidéo sans images ou contenant une consigne fixe, ce qui pourrait être utile pour de la compréhension orale. Au niveau des vidéos, le téléchargement d'une vidéo personnelle n'est pas non plus prévu ; celle-ci doit au préalable avoir été placée sur Youtube ou sur un autre serveur web pour pouvoir être intégrée.

Concernant les extensions disponibles, celles-ci se limitent à apporter des fonctionnalités sous forme de notifications, de possibilités d'ajout de formules mathématiques ou de graphiques. Il n'y a actuellement pas (encore) d'extension directement en lien avec une utilisation dans le monde de l'enseignement. On peut cependant citer l'extension Flubaroo, pour Google Spreadsheet, qui permet la correction automatisée de questions de type QCM.

Mais l'inconvénient majeur de cet outil réside dans sa limitation à corriger de manière automatisée. En effet, seule la correction directe des questions de type QCM est possible; pour les questions ouvertes, il est nécessaire de parcourir les réponses individuelles et de manuellement évaluer et attribuer des

points à chacun des items. Si l'on est au clair sur ce point, une solution élégante consiste à définir dans le quiz que la note sera attribuée après examen manuel, puis, une fois la correction manuelle et individuelle effectuée, d'envoyer un feedback complet à l'élève par email. Cette solution ne permettant cependant pas un retour rapide en classe, comme requis par mon projet, j'ai décidé de ne pas opter pour Google Forms.

[ProProfs \(www.proprofs.com\)](http://www.proprofs.com)

ProProfs est une plateforme de formation en ligne et d'évaluation, utilisée par d'illustres universités américaines et des sociétés de premier plan. Parmi les outils proposés, leur « Quiz Maker » permet la réalisation de quiz informatisés.

Avantages

ProProfs ne se cantonne pas à un outil de création de quiz. Il offre de nombreuses fonctionnalités complémentaires comme la mise en place d'un environnement d'entraînement, des flashcards, des jeux de logique, ainsi qu'une base de connaissance permettant la réutilisation de contenus ou questions.

Les types de questions proposés sont très riches : des choix multiples, cases à cocher, vrai/faux ou exercice de correspondance d'éléments entre eux sont disponibles pour les questions de type QCM. Pour les questions nécessitant la saisie de texte, des textes lacunaires ainsi que des essais sont à disposition. Il est également possible d'intégrer des fichiers audio et vidéo, et de permettre à l'apprenant de télécharger ses propres documents. Pour chaque type de question, il est possible d'inclure une image. L'intégration peut s'effectuer par téléchargement d'une image personnelle, via la bibliothèque intégrée de ProProfs ou une recherche Google.

L'interface peut quant à elle être totalement modifiée, soit par des thèmes prédéfinis, soit en modifiant librement les couleurs et les polices de caractère. Il est également possible de choisir la langue d'affichage du quiz et le site permet la contribution des créateurs de quiz pour la traduction d'éléments de l'interface. Un outil de prévisualisation puissant permet d'apprécier le résultat rapidement selon 3 formats (ordinateur / tablette / smartphone).

L'identification du preneur de quiz est prévue au sein de l'outil et différentes informations peuvent être récoltées : nom, email, id, téléphone, adresse, ou question personnelle.

L'affichage des questions est paramétrable ; celles-ci peuvent être affichées sur une page unique ou une question par page. La possibilité de naviguer ou non entre les questions est également un paramètre, tout comme le choix de vérifier l'ensemble des réponses données avant la soumission du quiz. Les réponses aux questions peuvent être révélées pendant le test ou seulement à la fin et les résultats sont totalement personnalisables ; en fonction du pourcentage de bonnes réponses, il est possible d'afficher un commentaire sous forme de réussi/échec, de lettre (A, B, C, etc), ou selon les besoins du créateur (par exemple des notes). Il est également possible de définir un temps limite pour la passation du quiz.

Au contraire de Google Forms, la correction des réponses fournies peut s'effectuer à la volée. Pour les réponses courtes, plusieurs réponses correctes peuvent être spécifiées par le créateur du quiz, en choisissant de prendre ou non en compte la casse. Un feedback détaillé peut en outre être donné pour chaque type de question.

Les résultats sont visualisables via une interface conviviale permettant de voir en un clin d'œil des informations telles que le temps moyen pour la complétion du quiz, le taux de réussite moyen ou par question, ainsi que les réponses par question. Les résultats peuvent être exportés au format Excel ou PDF.

Finalement, il est possible de spécifier la disponibilité du quiz. Celui-ci peut soit être à compléter avant une date butoire, soit accessible entre deux dates, soit ouvert ou fermé.

Inconvénients

L'inconvénient majeur de ProProfs réside dans le fait qu'il soit payant. Chaque outil de la plateforme est utilisable séparément mais nécessite un abonnement propre, ce qui peut être rapidement onéreux si l'on souhaite utiliser plusieurs fonctions. Dans le cas du Quiz Maker, cela représente un montant de 9 dollars par mois pour un maximum de 100 participants au quiz, ce qui est supportable.

Choix final :

Bien qu'offrant déjà de nombreuses fonctionnalités de quiz, Google Forms n'est à mes yeux pas encore suffisamment abouti, notamment en raison de la non-correction automatisée des questions ouvertes de type développement ou dans le cas d'une argumentation. Mon choix final s'est donc porté sur le service ProProfs pour mon dispositif, car il est actuellement meilleur que l'outil de Google.

A l'avenir, il faudra néanmoins garder un œil sur les améliorations apportées à Google Forms, notamment en matière de plugins destinés à l'enseignement.

Cependant, peu importe l'outil utilisé, pour qu'un quiz soit profitable aux élèves et aux enseignants, il est nécessaire qu'il soit fiable — dans des conditions similaires, un test devrait toujours donner les mêmes résultats (Harmer, 2015) — et bien conçu (Bravard, 2005). C'est le sujet du chapitre suivant, consacré à la conception des quiz.

1.3.3.4 Conception des quiz et définition des items

Bien que la littérature sur le sujet ne soit pas abondante, il est important de se pencher sur la problématique de la conception d'un quiz et de la définition des questions proposées.

Au niveau des questions, on distingue dans l'enseignement des langues secondes les items de test indirects et directs. Les premiers visent à tester la connaissance, par exemple de vocabulaire ou de grammaire, alors que les seconds demandent à l'élève de réaliser des tâches langagières directes, telles qu'écrire une lettre (Harmer, 2015). On distingue également les items de type *discrete point* et *integrative*. Le premier teste un élément à la fois, alors que le second regroupe plusieurs éléments et parfois plusieurs compétences dans un seul test (Harmer, 2015)..

En ce qui concerne les tests informatisés, seuls les items de type indirects peuvent être utilisés si l'on souhaite une correction automatique, car ce type de questions ne nécessite pas une appréciation ou pondération humaine. Harmer (2015) distingue les types d'items suivants :

- QCM : un des outils de test les plus populaires pour mesurer les connaissances en grammaire et vocabulaire, en particulier parce que la correction est aisée.
- Textes à trous: une phrase ou un texte est à compléter avec des mots ou des bouts de phrases. Dans les tests digitaux, les candidats ont souvent à glisser/déplacer l'item approprié depuis un container jusqu'à la bonne lacune.
- Réorientation : remettre les mots dans le bon ordre en dit beaucoup sur les capacités syntaxiques et lexico-grammaticales des élèves.
- Résumer: moyen de tester les capacités de compréhension des élèves et de retranscrire cette compréhension en mots.

Dans le cadre de mon dispositif, je vais me concentrer sur les 2 premiers types d'items car ce sont ceux utilisés par Roediger III et al. (2011).

Dans son étude de 2016, Fernandez révèle que les effets positifs de quiz peuvent disparaître dans certaines conditions. En effet, si les quiz proposés sont trop simples, les étudiants vont d'une part produire une activité cognitive et métacognitive inadaptée à la tâche finale demandée, et d'autre part surestimer la qualité de leur apprentissage. Il est donc primordial de concevoir des quiz de difficulté adaptée afin d'en maximiser les effets. Pour cela, j'ai décidé de varier les items de type *discrete point* et *integrative*, afin de proposer des questions selon différents niveaux de complexité.

Dans l'enseignement, de manière comparable à la taxonomie de Bloom, Baillat, De Ketele et al. (2008) définissent 6 niveaux de complexité des tâches : l'acquisition du vocabulaire s'inscrit dans le niveau le plus bas (niveau 1 : savoir restituer fidèlement un savoir). La grammaire et les règles appartiennent quant à elles au niveau 2 (savoir appliquer une démarche). Savoir utiliser une règle grammaticale en contexte est du niveau 3 (savoir appliquer une démarche dans un contexte), alors que des tâches complexes mobilisant les savoirs et les savoir-faire de base dans une situation problème appartiennent au niveau 4. Bien que le tableau comprenne 6 niveaux de complexité, les niveaux pouvant être testés à l'école secondaire se bornent au niveaux 1 à 4, les deux derniers nécessitant des compétences langagières très avancées.

Dans un quiz informatisé avec auto-correction, seuls les niveaux 1 à 3 pourront être testés car les niveaux supérieurs nécessitent une intervention humaine pour la correction. Il est donc intéressant de proposer dans les quiz une variété de questions appartenant aux 3 niveaux inférieurs.

Extraits	Démarches	Caractéristiques	Jargon
1	Savoir restituer fidèlement un savoir	Reconnaître un «objet», nommer, définir, énoncer...	« Connaissance » « Connaissance déclarative » « Savoir restituer »...
2	Savoir appliquer une démarche (algorithme, règle, procédure)	Appliquer une démarche apprise hors contexte et la consigne indique la démarche à appliquer	« Connaissance procédurale » « Application scolaire » « Savoir-faire de base »...
3	Savoir appliquer une démarche (algorithme, règle, procédure) dans un contexte	Appliquer une démarche apprise dans une situation, mais l'énoncé indique la démarche à appliquer	« Application habillée » « Savoir-faire de base en contexte »
4	Savoir mobiliser les savoirs et savoir-faire de base pertinents dans une situation problème	Mobiliser les savoirs et savoir-faire de base dans un contexte qui ne dit pas les savoirs et savoir-faire de base à mobiliser	« Savoir-faire complexe » « Tâche complexe » « Compétence » « Traitement de la situation »...
5	Comportements et attitudes qui sont passés dans l'habitude intériorisée	Des savoir-faire appris passent dans l'habitude intériorisée et sont des signes du système de valeurs de la personne	« Comportements » « Attitudes » « Posture » « Savoir-être »
6	Se mettre en projet, concevoir le projet, le réaliser, l'ajuster, le communiquer, le défendre	Identifier toute une série de ressources et de compétences pour réaliser un projet	« Projet » « Chef d'œuvre » « Savoir-devenir »

Figure 5: niveaux de complexité dans l'évaluation des apprentissages. Source : Baillat, De Ketele et al. (2008). Evaluer pour Former. DBS Pédagogie.

Concernant les énoncés des questions, Bravard (2005) définit un certain nombre de règles, parmi lesquelles :

- L'énoncé présente une seule notion soumise à la réflexion de l'apprenant. Beaucoup d'items, sans le vouloir, présentent plus d'un problème.
- L'énoncé précise l'objet de la question en utilisant un langage simple et clair. Il faut éviter de faire de vraies devinettes pour l'apprenant et plutôt faire en sorte que cela soit le plus compréhensible pour lui.
- L'énoncé est formulé, autant que possible, à la forme positive et éventuellement à la forme interrogative. En aucun cas il sera formulé à la forme négative. La présence de négations complique le traitement que l'apprenant doit opérer sur le texte. La négation n'est pas toujours exprimée par « ne pas » ou « non » ; elle est présente dans des mots tels que : retrait, absence, interdiction, cessation, nier, impossible, etc.
- L'énoncé doit être compatible avec la question.
- L'énoncé sera composé de tous les mots essentiels à sa compréhension. Il devra être le plus concis possible. De façon générale on peut dire que toute information requise à la solution d'un problème se retrouve dans l'énoncé plutôt que dans les choix de réponses.
- L'énoncé contient le mot à définir. Il pose la question tandis que dans les choix de réponses on inscrit les définitions possibles du mot demandé.
- L'énoncé précise l'objet de la question. Sans une telle précision le diagnostic risque d'être complètement erroné.

1.3.3.4.1 Les questions de type QCM

D'après Bravard (2005), le QCM est un type d'exercice très guidé qui a l'avantage de répondre assez bien aux contraintes liées à la programmation de quiz. Le terme QCM désigne tout exercice dans lequel l'apprenant répond à une question en choisissant la réponse parmi une liste de propositions. Ce type de questionnaire peut comprendre aussi bien les exercices d'appariement que les exercices lacunaires. Leclercq (1986) définit un QCM comme une série de questions auxquelles « l'étudiant répond en opérant une sélection (au moins) parmi plusieurs solutions proposées, chacune étant jugée (par le constructeur de l'épreuve et par un consensus entre spécialistes) correcte ou incorrecte indépendamment de l'apprenant qui doit y répondre ».

Une des critiques à l'égard des QCM concerne la place laissée au hasard, et donc à l'injustice qu'ils peuvent entraîner. En effet, comment différencier les « choix heureux par ignorance » des « choix corrects par compétence » (Leclercq, 1986). Une des parades envisagées consiste à augmenter le

nombre de réponses proposées, lorsque celui-ci n'est pas fixé. Une autre critique vise le fait qu'en choisissant une réponse parmi les alternatives proposées, l'apprenant n'a pas à concevoir sa propre réponse. Cet aspect peut être mis en relation avec les travaux de Croisile (2009) présentés plus haut ; si l'effort de récupération en mémoire est moindre, la trace mnésique sera faible et l'information plus sujette à l'oubli.

Un QCM comporte trois composantes, à savoir la consigne, l'énoncé et les propositions (Bravard, 2005).

La consigne décrit la question (une seule ou plusieurs propositions considérées comme correctes), le mode de réponse, les principes de notation et éventuellement le barème. La consigne doit être indiquée à l'apprenant en début de quiz ; elle reste valable pour toutes les mêmes questions de l'épreuve.

L'énoncé pose la situation de départ et pose la question. Il peut également être écrit sous la forme d'une phrase à compléter.

Les propositions suivent l'énoncé et sont en rapport avec celui-ci. L'apprenant doit distinguer la ou les réponses exactes parmi les propositions, en faisant abstraction des distracteurs possibles. La bonne réponse doit être exacte et peut, dans certains cas, se cantonner à être la meilleure réponse parmi celles proposées. Par distracteurs, on entend des réponses plausibles mais fausses. Elles peuvent être des erreurs fréquentes d'étudiants ou des idées erronées communément véhiculées.

Concernant les réponses possibles, Bravard (2005) énonce quelques règles à respecter pour obtenir des quiz de qualité:

- La bonne réponse, tout au long du test, varie de place de façon aléatoire. Il est donc judicieux d'utiliser un outil faisant varier l'ordre des réponses possibles.
- Les choix de réponses sont homogènes dans leur contenu, leur forme et leur structure grammaticale. C'est sans doute une des règles les plus difficiles à respecter. S'il est aisé de trouver la bonne réponse, il est par ailleurs difficile de choisir des distracteurs faux mais plausibles, et qui respectent toutes les règles de la construction des choix de réponses.
- Tous les distracteurs sont plausibles mais faux. Il est préférable de mesurer certaines notions par d'autres formes de questions plutôt que d'écrire des choix de réponses non plausibles et qui n'ont rien à voir avec l'énoncé, uniquement pour avoir un nombre égal de choix de réponses dans tout l'examen.
- Il est préférable de ne pas répéter un même mot dans les choix de réponses. Pour éviter qu'un

mot ne se répète dans chacun des choix de réponses, on l'intègre à l'énoncé. Il est à noter que, pour ces cas, le mode affirmatif est plus approprié que le mode interrogatif.

- Eviter que certains mots compris dans l'énoncé ne se répètent dans un choix de réponses et conduisent ainsi à la bonne réponse.

1.3.3.4.2 Les questions de type réponses courtes

Chan (2009) définit les questions à réponses courtes comme des questions ouvertes demandant aux étudiants de formuler par eux-mêmes une réponse. Ce type de question est similaire aux questions de type « essai » mais la longueur de la réponse est limitée à une certaine taille, allant d'un mot à plusieurs lignes.

Les questions de type réponse courte ne possèdent pas de structure générique. Les types de réponses attendues peuvent ainsi grandement varier : phrase complète, remplissage d'une lacune, courte description, ou encore diagrammes avec explications (Chan, 2009).

Au contraire de questions de type « essai », mais de manière similaire aux questions à choix multiple, les questions de type réponse courte ne testent que des domaines inférieurs de connaissance, comme des connaissances factuelles. Cela est dû au format de réponse attendu, qui ne permettent pas d'élaborer des réponses complexes (Chan, 2009).

Concernant les avantages de ce type de questions, Chan (2009) en évoque 3 principaux : elles sont relativement aisées à corriger et la correction peut être effectuée par des correcteurs différents sans impacter le résultat. En comparaison d'autres méthodes d'évaluation, elles sont aisées à mettre en place. Enfin, au contraire des questions de type QCM, il n'est pas possible de deviner la réponse ; les élèves doivent fournir une réponse d'eux-mêmes.

Les questions de type réponses courtes nécessitent néanmoins un soin important dans la définition de la question ; du fait que la réponse est libre, un manquement dans la formulation de la question peut entraîner des difficultés au niveau de l'élève de ce qui est attendu de lui. Un manquement à ce niveau peut également entraîner des soucis de correction, des réponses inattendues étant formulées. De plus le temps nécessaire à la réponse de questions de ce type est plus conséquent que pour des questions de type QCM (Chan, 2009).

Dans le but de faciliter la création de questions à choix ouvert, Chan (2009) fait les recommandations suivantes :

- créer des questions qui mesurent les objectifs d'apprentissage de manière appropriée.
- exprimer les questions avec des termes clairs et un langage approprié pour les élèves.
- faire en sorte qu'il n'y ait qu'une réponse correcte à chaque question.
- s'assurer que la question exprime clairement le type de réponse attendu (mot seul, courte phrase, description).
- écrire les instructions clairement de manière à spécifier la connaissance voulue et le type de réponse attendu.
- être préparé à accepter d'autres réponses, également acceptables, par rapport aux réponses prévues.
- préparer une feuille de correction structurée, indiquant les réponses possibles et les points alloués à chaque question.

1.4 Questions et objectifs de recherche

La présente étude cherche à répondre à deux questions de recherche :

1. Est-ce que les résultats de Roediger III et al. 2011 s'appliquent à une classe de langue seconde du secondaire 1 ?
2. Quelle est la forme à donner aux tests de rappel pour obtenir les meilleurs résultats ?

L'étude de Roediger III et al. s'étant focalisée sur un cours de sciences sociales, il est intéressant de voir si des résultats similaires sont obtenus dans le cadre d'une classe de langue seconde du secondaire 1. D'après les auteurs d'une autre étude (McDaniel et al., 2013), outre les domaines de la physique et de la médecine, les langues pourraient également bénéficier du *testing effect*. Si une amélioration de la rétention d'information après des tests est observée, il s'agit également de voir si le test de réponses courtes donne les meilleurs résultats.

Hypothèse : les résultats obtenus avec des élèves d'âge similaire dans une matière différente se vérifient dans un cours de langue. A savoir :

- les quiz à réponses courtes obtiennent des résultats plus élevés comparés aux questionnaires à choix multiples
- la matière ne bénéficiant d'aucun dispositif obtient le plus bas pourcentage de bonnes réponses à l'évaluation sommative.

La première partie de mon travail cherche à vérifier mon hypothèse que les résultats de Roediger III et al. (2011) s'appliquent à une classe de langue du secondaire 1.

La seconde partie de mon étude vise à répondre à la deuxième question de recherche. A mes yeux, les dispositifs utilisés dans les études précédentes sur le *testing effect* comportent un défaut majeur : ils ne sont pas applicables pour un enseignant « lambda » dans sa pratique quotidienne. La seconde partie de mon travail sera donc l'occasion de définir un outil approprié et une méthodologie de tests pour palier ce problème. Il s'agira, via une étude qualitative, de définir notamment :

- L'outil à utiliser pour les tests (technologie, interface)
- Le contenu des tests
- La forme du feedback apporté
- La fréquence des tests

Rapport-Gratuit.com

2. Première partie

2.1 Méthodologie

La première partie de mon travail cherche à vérifier mon hypothèse que les résultats de Roediger III et al. (2011) s'appliquent à une classe de langue seconde du secondaire 1, à savoir que lors d'une évaluation finale, les questions préalablement testées au format réponses courtes obtiennent de meilleurs résultats par rapport aux questionnaires à choix multiples, et que la matière qui n'a bénéficié d'aucun dispositif obtient le moins bon résultat à l'évaluation sommative.

2.1.1 Fondements méthodologiques

La démarche méthodologique pour cette partie consiste à réaliser une recherche expérimentale. Selon Myers & Hanson (2007) ce type de recherche vise à « vérifier une prédiction, appelée hypothèse [...] qui vise à montrer que certains types d'événements sont prévisibles dans certaines situations spécifiques ». Dans mon cas, l'hypothèse à vérifier est que les résultats de Roediger III et al. (2011) obtenus dans un cours de sciences sociales s'appliquent également à un cours de langue seconde. Dans leur étude de 2011, Roediger III et al. ont mené 3 expériences distinctes : la première consistait à voir si des tests amélioraient la rétention d'information par rapport à des conditions de contrôle impliquant l'absence de tests ou la relecture de la matière. Cette recherche a été menée avec des questions de type QCM et rappel libre – consistant à écrire sur une feuille de papier tout ce qu'on se rappelle d'un chapitre. La seconde expérience avait pour but de déterminer si les tests étaient plus bénéfiques que de multiples révisions de la matière. Ici, toutes les questions proposées étaient de type QCM. Enfin, la 3^e expérience avait pour but de voir si les bénéfices de la récupération répétée d'informations en mémoire via des tests s'étendaient à des questions de type réponses courtes. Mon étude consiste en un mélange entre les expériences 1 et 3 de Roediger III et al. (2011), à savoir la vérification de l'existence d'un *testing effect* mais avec des questions de type QCM et réponses courtes.

La méthode utilisée est l'observation de la variation des résultats d'élèves d'une classe après les avoir exposés à un dispositif censé valider ou invalider mon hypothèse, selon le design pré-test/post-test (Karsenti & Savoie-Zajc, 2004).

2.1.2 Nature du corpus

2.1.2.1 Choix de la branche d'enseignement

Ayant suivi la didactique d'allemand en 1^{ère} année et suivant actuellement celles de français et d'anglais, je me sens plus à l'aise de mener l'étude dans la branche où je suis déjà formé et où je bénéficie d'une année d'expérience professionnelle. De plus, j'aimerais éviter d'introduire un biais dans mon étude, qui serait le fait que je ne suis pas formé dans la branche d'étude. Pour ces raisons, j'ai choisi l'allemand comme branche pour mon étude.

2.1.2.2 Population

Les participants sont les 20 élèves (11 garçons et 9 filles) d'une classe de 9^e Harmos du Centre Scolaire Le Locle (CSLL), dans laquelle j'enseigne l'allemand. L'âge des jeunes se situe entre 12 et 15 ans. Ces élèves possèdent un niveau d'allemand élémentaire assimilable au niveau A1 du CEFR⁷. Leurs parents ont été informés de l'étude par une circulaire. Vu que l'évaluation finale du dispositif consiste en une évaluation sommative comptant dans la moyenne annuelle, j'ai choisi de ne pas scinder la classe en deux groupes distincts (groupe expérimental et groupe témoin) car il ne serait éthiquement pas correct de mettre en place un dispositif censé améliorer les résultats pour une partie de la classe uniquement. Au lieu de cela, le groupe classe au complet est passé successivement dans les diverses situations expérimentales: les mêmes élèves ont été testés plusieurs fois, dans des contextes différents ; il s'agit ainsi d'un plan expérimental à groupes appariés (Sockeel & Anceaux, 2008).

2.1.2.3 Contenu didactique

La matière utilisée pour l'étude a été celle prévue par le plan d'étude pour le début du deuxième semestre de la 9^e année Harmos, en particulier l'unité 12 du livre d'allemand Geni@! A1, qui traite de l'orientation dans une école. Dans ce chapitre, les objectifs de compréhension de l'écrit et de l'oral étaient:

1. Comprendre de manière globale et sélective :
 - a. le nom des différentes salles d'une école et ce que l'on peut y faire
 - b. à quel étage se situe une salle
2. Comprendre de manière globale et sélective :
 - a. où se trouvent les différents objets d'une salle de classe
 - b. comment les différents objets sont positionnés

⁷ Cadre européen commun de référence pour les langues



3. Comprendre de manière globale et sélective quelqu'un qui demande et décrit :
 - a. où se trouve une salle
 - b. le chemin pour y parvenir
4. Comprendre les principales informations d'un texte ayant pour thème l'école (qui, quand, quoi, où, ...)

Les objectifs de production de l'écrit et de l'oral sont :

1. Indiquer :
 - a. le nom des différentes salles d'une école et ce que l'on peut y faire
 - b. à quel étage se situe une salle
2. Demander et indiquer :
 - a. où se trouvent les différents objets d'une salle de classe
 - b. comment les différents objets sont positionnés
3. Demander et indiquer :
 - a. où se trouve une salle
 - b. le chemin pour y parvenir

Au niveau du fonctionnement de la langue, l'élève doit être capable :

- de conjuguer et d'utiliser les verbes de positionnement (*sein, stehen, liegen, sitzen, hängen*)
- d'utiliser les prépositions de lieu au datif (*an, auf, hinter, in, neben, über, unter, vor, zwischen*)
- d'utiliser les nombres cardinaux pour indiquer les différents étages
- d'utiliser le vocabulaire en lien avec l'unité 12

Ce chapitre a été traité sur une durée de 5 semaines, à raison d'une dotation horaire de 4 périodes (3 périodes en plenum, 1 période en demi-groupe).

2.1.2.4 Procédure et protocole de l'expérience

Sur un laps de temps de 5 semaines (soit le temps moyen nécessaire pour couvrir une unité du livre d'allemand de 9^e Harnos), des tests ont été administrés chaque semaine, sous forme de quiz électroniques de 15 minutes. Ces quiz ont été passés lors des séances en plenum du jeudi ou vendredi, selon l'avancement dans la matière de la semaine. Ils étaient composés de questions sous forme de QCM et de réponses courtes à donner, et couvraient la moitié de la matière vue pendant la semaine. Le nombre de questions variait entre 9 et 16, selon la semaine. Pour chaque réponse, un feedback était donné à l'élève et son résultat au quiz lui était communiqué en fin de test. Un ultime quiz de révision a été passé par les élèves 24 heures avant l'évaluation sommative, qui comptait dans la moyenne

annuelle des élèves et clôturait la séquence. Ce quiz comprenait 53 questions, soit l'ensemble des questions présentes dans les quiz initiaux, mais affichées aléatoirement. Les alternatives pour les questions de type QCM étaient également réordonnées au hasard.

Une évaluation sommative de validation au format papier a été passée par les élèves en fin de chapitre. Celle-ci comportait 39 questions, dont les deux tiers étaient des questions testées dans les quiz précédents (12 questions au format QCM, 12 autres sous forme de réponses courtes), et un tiers des questions sur la matière vue en cours, mais non testée en classe. Afin de garantir une uniformité des réponses données à l'évaluation finale, toutes les questions ont été présentées au format réponses courtes.

Pour des raisons de répartition horaire, le quiz de révision a été passé le jeudi de la dernière semaine de l'étude, et l'évaluation sommative le vendredi, afin de respecter les 24 heures de délai entre les deux tests.

Comme communiqué aux élèves en début d'étude, les résultats des tests intermédiaires n'ont pas été pris en compte dans la note finale. Il s'agissait juste d'un bonus en vue de la préparation du test final. Ce choix a été effectué après analyse des résultats des élèves à un quiz testé en amont de l'étude, qui se sont avérés très mauvais. De plus, afin d'éviter l'achat de matériel dédié, comme des boîtiers de vote, ou un déplacement en salle d'informatique lors de chaque quiz, j'ai décidé d'utiliser les smartphones des élèves pour la passation des tests.

Une fois l'expérimentation en classe réalisée, les résultats de l'évaluation sommative ont été analysés afin de savoir si les éléments préalablement testés obtenaient de meilleurs scores au test final que la matière non testée. Pour ce faire, j'ai divisé les résultats obtenus en trois performances : une sur le contenu testé au préalable au format QCM (performance A), une seconde sur le contenu testé en mode réponses courtes (performance B), et la dernière sur le contenu non testé (performance C). Par la suite, j'ai utilisé la technique de l'analyse descriptive afin de :

- voir s'il y a une progression avérée entre les questions présentes dans les quiz et la matière non testée.
- décrire le nombre de sujets qui ont progressé, ceux qui ont stagné, et ceux qui ont régressé.
- voir si le dispositif favorise tous les élèves, ou seulement les bons, ou les moins bons.

La variable principale de mon dispositif est le fait que la matière ait été testée ou non en amont de l'évaluation sommative. Cette variable a permis de répondre à ma première question de recherche. Une

seconde variable est le type de questions proposées (QCM ou réponses courtes). Cette variable me permettra de voir si les résultats de Roediger et al. (2011) s'appliquent également pour ma classe d'allemand, à savoir que les contenus préalablement testés via des réponses courtes obtiennent des résultats plus élevés lors de l'évaluation sommative. Enfin, j'utiliserai une variable qualitative qui me permettra de classer les élèves selon leur niveau, sur la base de leurs notes annuelles. Cette variable me permettra de voir à quel profil d'élèves profite le plus mon dispositif.

Afin de vérifier si les résultats obtenus étaient statistiquement significatifs, je leur ai appliqué différents outils statistiques. Si la valeur p de ces tests était inférieure à 5%, indiquant un taux de fiabilité de 95%, les résultats obtenus pouvaient être considérés comme statistiquement significatifs.

Les 4 quiz initiaux ainsi que l'évaluation sommative sont disponibles en annexe (annexes 1 à 5).

2.1.2.5 Différences par rapport à l'étude de base

Par rapport à l'étude de Roediger III et al. (2011), mon étude diffère sur certains points et ne peut donc en ce sens pas être considérée comme une réplique. Les différences se rapportent principalement au mode de passation des tests, à la nature des participants et aux tests de rétention à long terme.

Concernant le mode de passation des tests, le fait qu'ils soient administrés par l'enseignant et non par un assistant de recherche représente la première différence. L'idée maîtresse étant d'offrir un dispositif léger à mettre en place par l'enseignant, il paraissait normal qu'il puisse le faire de manière autonome, sans aide extérieure. La seconde différence concerne l'abandon du pré-test initial faisant suite à la lecture autonome d'un texte à la maison. En effet, alors que l'enseignement testé était prodigué dans la langue maternelle des élèves dans les autres études – sciences sociales pour l'étude de Roediger & al. (2011), sciences pour les études de McDaniel & al. (2011) et McDaniel & al. (2013) – l'allemand est pour tous les élèves de la classe testée une langue seconde et leur connaissance de la langue est trop restreinte pour leur permettre de comprendre un texte sur un nouveau thème de manière autonome à la maison. J'ai donc décidé de me concentrer sur les tests en fin de séquence d'enseignement, ainsi que le test de révision 24 heures avant l'évaluation sommative finale.

Pour des raisons de gain de temps pour l'enseignant et du fait de la répartition des leçons d'allemand sur la semaine (quatre périodes réparties en trois leçons en plenum sur trois jours différents, ainsi qu'une période en demi-groupe), les tests systématiques après chaque leçon ont été remplacés par un quiz hebdomadaire, couvrant la matière de la semaine. Pour des raisons pratiques et financières,

l'utilisation des smartphones des élèves a été préférée au mode de passation par boîtiers. Les questions n'ont pas été lues à haute voix à toute la classe, mais affichées sur les smartphones au fur et à mesure que les élèves avançaient dans le quiz. Enfin, le feedback a également été donné sous une autre forme : au contraire d'un feedback immédiat affiché sous la forme d'une coche verte à côté de la bonne réponse au beamer, le feedback a été différé à la fin du quiz et communiqué en même temps que le résultat final. Les élèves avançant à leur rythme dans le quiz, le feedback donné oralement après chaque question a également été abandonné.

L'évaluation sommative finale comportait également une différence mineure par rapport à l'étude de base. Au contraire d'alterner les questions proposées (selon le schéma suivant : 1 question précédemment testée au format QCM, 1 testée sous forme réponse courte, et 1 question non testée) comme prévu dans l'étude de base, j'ai décidé de m'affranchir de ce critère et ainsi de proposer des questions regroupées de manière logique par thèmes. Ce choix a été effectué pour des raisons éthiques, afin d'éviter que les élèves ne soient déstabilisés par un changement constant de thèmes dans les questions proposées.

Au sujet des participants à l'étude, l'âge des élèves est légèrement supérieur – 12 à 15 ans contre 11 à 12 – et leur nombre est restreint à 20, soit le nombre d'élèves de ma classe d'allemand.

Finalement, contrairement à l'étude initiale où des tests similaires à l'évaluation finale ont été administrés à la fin du semestre et à la fin de l'année scolaire, mon étude s'est cantonnée à l'évaluation finale de fin de chapitre et ne prend pas en compte cet aspect de rétention à long terme.

Ces différences ont été dictées par des raisons pratiques – je n'ai qu'une seule classe en responsabilité d'allemand, ce qui limite le nombre de participants possibles – et temporelles, l'étude initiale ayant été menée sur trois semestres, ce qui n'était pas possible dans le cadre du présent travail.

2.1.3 Méthode de collecte des données

Une fois l'expérimentation en classe terminée, les évaluations sommatives ont été corrigées et les résultats transposés en points, puis placés dans une table Excel. Pour chaque question, un maximum de 1 point était possible et selon les erreurs commises, les élèves étaient crédités de 1 point, 0.75, 0.5, 0.25 ou 0 point par question.

Au cours de l'expérimentation, certains élèves ont manqué pour diverses raisons un ou plusieurs quiz et ont donc été écartés de l'échantillon de données. Au final, sur les 20 élèves participants, seuls les résultats de 11 élèves ont pu être pris en compte. Le tableau ci-après résume les raisons qui m'ont fait écarter 9 élèves :

Elève	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Review Quiz	Evaluation sommative
E2	Présent	Absent	Présent	Présent	Présent	Présent
E7	Absent	Présent	Présent	Présent	Présent	Présent
E8	Présent	Présent	Absent	Présent	Absent	Absent
E10	Présent	Absent	Présent	Présent	Présent	Présent
E13	Absent	Présent	Présent	Absent	Absent	Absent
E14	Présent	Présent	Présent	Absent	Présent	Présent
E15	Présent	Présent	Absent	Présent	Présent	Présent
E19	Présent	Présent	Absent	Absent	Absent	Absent
E20	Présent	Pas pu terminer (résultat de 0%)	Présent	Présent	Présent	Présent

Tableau 1: absences de certains élèves aux quiz

2.2 Analyse

2.2.1 Ecarts de performance entre les 3 types de questions

A partir des données récoltées, j'ai calculé la moyenne des résultats de chaque élève pour les 3 types de questions proposées, à savoir les questions préalablement testées au format QCM, au format réponses courtes, ou les questions non testées au préalable.

	EI. 1	EI. 2	EI. 3	EI. 4	EI. 5	EI. 6	EI. 7	EI. 8	EI. 9	EI. 10	EI. 11
Résultat aux questions préalablement testées au format QCM	63.46%	69.23%	84.62%	19.23%	78.85%	94.23%	67.31%	50.00%	65.38%	34.62%	98.08%
Résultat aux questions préalablement testées au format choix ouvert	61.54%	71.15%	90.38%	44.23%	73.08%	82.69%	69.23%	65.38%	75.00%	61.54%	92.31%
Résultat aux questions non testées	48.08%	59.62%	75.00%	15.38%	67.31%	80.77%	57.69%	42.31%	51.92%	42.31%	78.85%

Tableau 2: moyennes des résultats des élèves, pour les 3 types de questions

Par la suite, j'ai calculé la moyenne des moyennes de tous les élèves pour chaque type de question, afin de pouvoir analyser les résultats selon le type de question. On constate que la matière non testée obtient la performance la moins élevée (56.29%), contre 65.91% pour les questions testées au format

QCM. La performance pour les questions testées au format réponses courtes est la plus élevée avec plus de 71% de bonnes réponses.

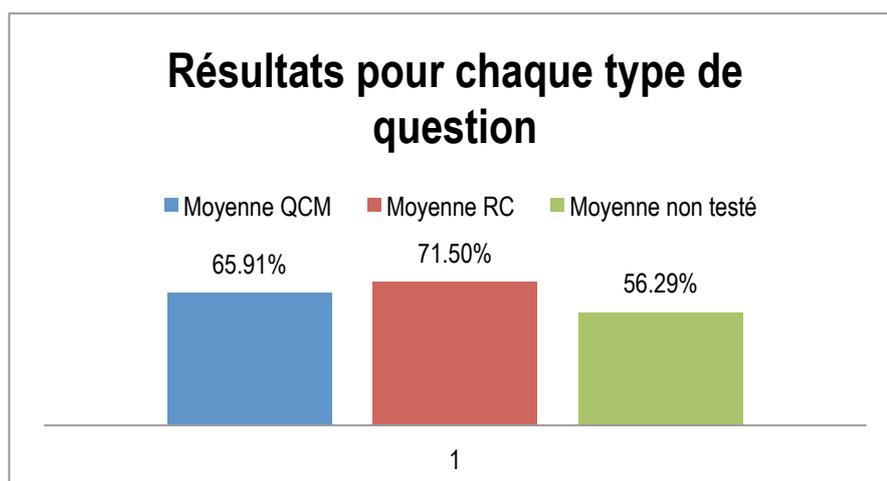


Tableau 3: résultats pour chaque type de question

Afin de vérifier si les différences de performance observées étaient statistiquement significatives, j'ai utilisé le test de Friedman. Celui-ci s'applique pour des études qualitatives sur plus de deux groupes appariés, lorsque la variable de réponse est également qualitative⁸. Le test a révélé une valeur $p=0.000462$ (soit 0.0462%), largement en-dessous du seuil de tolérance de 5%. Cet outil a donc permis d'affirmer que les différences observées entre les trois moyennes sont statistiquement significatives.

Ces résultats confirment les hypothèses formulées en début de recherche, à savoir que les questions de type réponses courtes obtiennent des résultats plus élevés comparés aux questionnaires à choix multiples, et que la matière n'ayant bénéficié d'aucun dispositif obtient le pourcentage de bonnes réponses le moins élevé lors de l'évaluation sommative. Je peux donc répondre à ma première question de recherche par l'affirmative : les effets observés par Roediger III et al. (2011) se reproduisent dans une classe d'allemand du secondaire 1.

2.2.1.1 Analyse des moyennes en ligne pour chaque question

Au vu des résultats obtenus pour chaque type de question proposé, j'ai souhaité savoir si les moyennes des résultats des élèves pour chaque question étaient cohérentes, et ainsi voir si les questions proposées étaient adaptées. Ces moyennes sont présentées dans le tableau ci-après :

⁸ Selon l'outil BiostaTGV, accessible en ligne à l'adresse <http://marne.u707.jussieu.fr/biostatgv/?module=tests>

Question	Mode de test initial	Moyenne des réponses correctes
2	Non testé	0.33
4	Non testé	0.64
8	Non testé	0.75
10	Non testé	0.75
12	Non testé	0.52
13	Non testé	0.95
15	Non testé	0.70
20	Non testé	0.73
23	Non testé	0.45
28	Non testé	0.23
29	Non testé	0.14
33	Non testé	0.50
34	Non testé	0.64
7	QCM	0.70
11	QCM	0.73
14	QCM	0.73
18	QCM	0.55
19	QCM	0.82
24	QCM	0.64
26	QCM	0.32
27	QCM	0.50
31	QCM	0.80
32	QCM	0.55
35	QCM	0.95
36	QCM	0.36
38	QCM	0.93
1	Choix ouvert	0.73
3	Choix ouvert	0.86
5	Choix ouvert	0.93
6	Choix ouvert	0.89
9	Choix ouvert	0.95
16	Choix ouvert	0.84
17	Choix ouvert	0.82
21	Choix ouvert	0.50
22	Choix ouvert	0.45
25	Choix ouvert	0.18
30	Choix ouvert	0.82
37	Choix ouvert	0.59
39	Choix ouvert	0.73

Table 1: moyennes des résultats par question

L'analyse en ligne des moyennes a mis en évidence certaines irrégularités dans les résultats, sous la forme de résultats largement inférieurs (en rouge dans le tableau) ou largement supérieurs à la moyenne (en vert) pour le type de question. Pour les questions non testées au préalable, on remarque que 3 questions (2, 28 et 29) ont obtenu des résultats inférieurs à 35% de réussite, alors que 4 questions (8, 10, 13, 20) ont obtenu des moyennes largement supérieures à la performance du type de question.

Pour les questions de types QCM, la question 26 a obtenu une moyenne inférieure à 35%, alors que les questions 18, 27, 32 et 36 sont légèrement inférieures ou égales à 55%.

Enfin, la 1 question (25) a obtenu un résultat inférieur à 20% pour les questions de type choix ouvert, alors que 2 autres questions (21, 22) ont obtenu une moyenne inférieure ou égale à 50%.

Pour les questions 25 à 29, la consultation de l'évaluation sommative a permis de mettre en évidence une piste pouvant expliquer ces mauvais résultats. En effet, toutes ces questions traitaient des verbes de modalité, qui n'appartenaient pas à la matière de l'unité 12 mais à celle de l'unité 11. Bien que 2 verbes sur 4 aient été testés dans le premier quiz, il est fort probable que l'éloignement temporel entre cette matière et celle de l'unité 12 soit l'explication de ces résultats. Pour les autres questions dont les résultats comportent d'importantes fluctuations, aucune piste d'explication ne peut être avancée.

En considérant que les questions comportant des résultats largement déviants par rapport à la moyenne du type de question pourraient être des biais éventuels, j'ai retiré les questions en cause et effectué à nouveau le calcul de la moyenne des résultats, avec les résultats suivants : les questions non testées progressent peu avec un pourcentage de 57.58%, alors que les questions testées au format QCM bondissent à 76.01%. En haut du tableau, les questions de type réponses courtes affichent une moyenne impressionnante de 81.59%. Ces résultats sont en adéquation avec ceux obtenus plus hauts, mais indiquent que le choix des questions proposées peut avoir une incidence non négligeable sur les résultats.

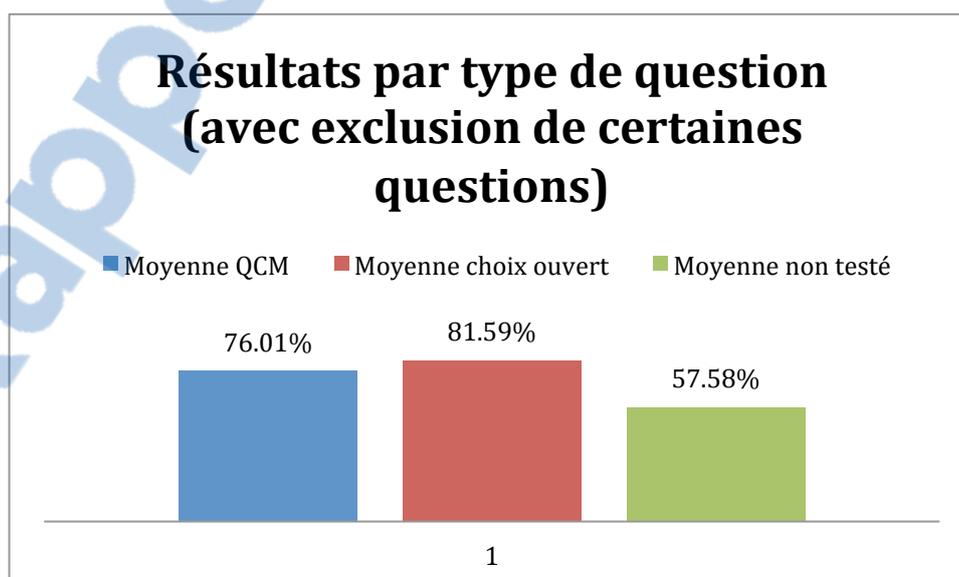


Table 2: résultats par type de question, après exclusion

2.2.2 Incidence du dispositif sur la performance des élèves

Afin de pouvoir analyser l'incidence du dispositif mis en place sur la performance des élèves, j'ai cherché à décrire le nombre d'élèves qui ont progressé, ceux qui ont stagné et ceux qui ont régressé entre le début et la fin de mon étude. Pour ce faire, j'ai comparé leur moyenne annuelle au 10 février, soit avant le début de l'expérimentation en classe, avec la note obtenue lors de l'évaluation sommative finale.

Moyenne au 10.02	Note évaluation sommative	Différence
5.95	6	0.84%
4.77	5	4.82%
5.00	4.5	-10.00%
4.59	3.5	-23.75%
5.27	5.5	4.36%
3.36	2	-40.48%
4.36	4.5	3.21%
5.32	5.5	3.38%
4.45	4.5	1.12%
3.36	3.5	4.17%
4.00	4	0.00%

Table 3: comparaison de la moyenne annuelle et des notes de l'évaluation sommative

Sur les 11 élèves dont les résultats ont pu être pris en compte dans mon étude, il apparaît que le dispositif a été positif pour 7 d'entre eux (63.63%), alors que 3 élèves ont obtenu une note inférieure à leur moyenne annuelle lors de l'évaluation sommative (27.27%). Les résultats d'un élève n'ont apparemment pas été influencés par le dispositif (9.09%). Si l'on s'intéresse aux différences, il apparaît que les notes des élèves ayant progressé ont augmenté en moyenne de 3.13%, alors que celles des élèves ayant régressé ont baissé de 24.74% en moyenne. Il semblerait donc que le dispositif ait été favorable à 63.63% des élèves, mais dans de moindres proportions, alors que 27.27% des élèves a enregistré des résultats en deçà de près de 25% par rapport à leur moyenne annuelle.

L'application du test de Student sur ces données, en prenant comme première variable la moyenne annuelle et comme seconde variable la note à l'évaluation sommative a cependant révélé une valeur de $p=0.32288$ (32.288%), soit largement supérieure à l'intervalle de confiance de 5%. Les différences obtenues ne peuvent ainsi pas être considérées comme statistiquement significatives.

2.2.3 Ecart de performance selon les catégories d'élèves

Une autre question évoquée dans la partie méthodologique du présent travail visait à voir si le dispositif favorisait tous les élèves ou seulement une partie (bons ou moins bons élèves). Les observations effectuées au point 2.2.2 permettent déjà d'affirmer que le dispositif ne favorise pas tous les élèves. Cependant, pour définir quel type d'élève bénéficie du dispositif, il s'agit de calculer la médiane des moyennes des élèves de la classe, puis de les répartir en deux groupes de par et d'autre de cette médiane. Le groupe des bons élèves regroupe les élèves ayant une moyenne supérieure à la médiane, alors que les moins bons sont les élèves dont la moyenne est inférieure à la médiane.

<i>Médiane des moyennes = 4.59</i>					
Groupe 1 (inférieur à la médiane)			Groupe 2 (supérieur à la médiane)		
<i>Moyenne</i>	<i>Note</i>	<i>Progression</i>	<i>Moyenne</i>	<i>Note</i>	<i>Progression</i>
3.36	2	<	5.95	6	>
3.36	3.5	>	5.32	5.5	>
4.00	4	=	5.27	5.5	>
4.36	4.5	>	5.00	4.5	<
4.45	4.5	>	4.77	5	>
Elève dont la moyenne correspond à la médiane					
<i>Moyenne</i>	<i>Note</i>	<i>Progression</i>			
4.59	3.5	<			

Table 4: répartition des élèves en deux groupes, selon leur moyenne annuelle

Parmi les élèves ayant bénéficié du dispositif, on remarque que 3 élèves appartenant au groupe inférieur (60% du groupe 1) ont bénéficié du dispositif, contre 4 élèves (80% du groupe 2) du groupe supérieur. Le dispositif semble ainsi favoriser les « bons » élèves, à savoir ceux dont la moyenne annuelle est supérieure à la médiane. L'élève dont la moyenne annuelle correspond à la médiane à quant à lui été pénalisé par le dispositif. Si l'on s'intéresse à la valeur de la progression, la moyenne des sujets du groupe 1 ayant progressé est de 2.83%, contre 3.35% au groupe 2.

Ces résultats sont à relativiser car l'application répétée du test de Friedman sur les 2 groupes de 5 sujets a révélé des valeurs de $p=0.31173$ (31.173%) pour le groupe 1 et $p=0.1797$ (17.97%) pour le groupe 2, soit largement au-dessus du seuil de tolérance de 5%. Les résultats obtenus ne sont donc pas statistiquement significatifs.

2.2.4 Discussion

Le dispositif conçu pour la recherche expérimentale a l'avantage de s'intégrer à une pratique d'enseignement existante, sans chambouler le quotidien de l'enseignant; la matière utilisée correspond à celle des chapitres étudiés, alors que l'évaluation finale du dispositif est l'évaluation de validation du chapitre. Le coût hebdomadaire des tests est également moindre, les tests s'effectuant en 15 minutes à raison d'une fois par semaine. Enfin, comme les tests sont passés par les élèves via leur téléphone portable, ils ne nécessitent pas d'acquisition de matériel ou de déplacement en salle d'informatique.

Ce dispositif comporte cependant des limites. D'abord, seule la partie écrite de l'enseignement langagier est testée et évaluée. On ne tient pas compte de l'écoute et de la production orale. Le nombre limité de participants est également limitatif car les résultats ne sont pas représentatifs à large échelle. Il serait souhaitable de mener cette étude sur plusieurs centaines d'élèves pour obtenir des résultats généralisables. Enfin, bien que prouvant l'existence d'un *testing effect* sur du contenu d'un cours d'allemand au secondaire 1, les résultats montrent que le dispositif déployé n'est pas bénéfique pour l'ensemble des élèves. Afin de tenter de comprendre ce qui n'a pas fonctionné au sein de mon dispositif pour certains élèves, j'ai sondé les représentations des élèves à l'aide de focus groups. Ce point est traité en détail dans la deuxième partie de la présente étude. Sur la base de ces retours en focus groups, ainsi que sur mes réflexions personnelles, les pistes d'explication suivantes peuvent être avancées:

- les résultats aux tests hebdomadaires ne comptant pas dans la moyenne, cela peut engendrer un engagement cognitif moindre.
- le format du test sommatif, très différent des tests effectués depuis le début de l'année, peut avoir déstabilisé certains élèves, notamment ceux rencontrant des difficultés en allemand.
- le mode de feedback choisi – différé à la fin du test, sauf pour le quiz 4 où il a été immédiat après chaque question – n'a peut-être pas permis à certains élèves de prendre conscience de leurs erreurs. En effet, le fait d'avoir tous les résultats affichés sur une page récapitulative en fin de quiz n'est peut-être pas judicieux, surtout lorsque les erreurs sont nombreuses. Il serait intéressant, dans le cadre d'une recherche complémentaire, de s'intéresser aux questions du quiz 4 afin de voir si elles obtiennent de meilleurs résultats lors du quiz de révision et lors de l'évaluation finale.
- le placement du test – en fin de leçon le jeudi ou le vendredi – n'était peut-être pas optimal. Une fois le test terminé, les élèves n'avaient probablement pas envie de prendre encore le temps de lire le feedback en détail, mais plutôt de quitter l'école, vu qu'ils avaient déjà leurs résultats.

Les résultats de la classe test sur toute la durée du dispositif (annexe 6), les résultats des élèves à l'évaluation sommative (annexe 7), les calculs pour l'analyse de la première partie de l'étude (annexe 8), ainsi que les scores des questions de l'évaluation finale (annexe 9) sont disponibles en annexe.

3. Deuxième partie

3.1 Méthodologie

3.1.1 Fondements méthodologiques

La seconde partie de mon étude vise à répondre à la deuxième question de recherche. Il s'agit là d'une étude qualitative menée dans le but de définir notamment l'outil à utiliser pour les tests, son interface, le contenu proposé, ainsi que la fréquence des tests. L'objectif étant de définir un dispositif permettant à tout enseignant de réaliser des tests de rappel de manière autonome, au contraire des études préalables sur le *testing effect*, où un assistant était présent en classe.

J'ai choisi d'effectuer une recherche pragmatique sous la forme d'une recherche-action. Selon Dolbec & Clément (2004), ce type de recherche vise à intégrer l'action au processus de recherche afin que le changement visé se produise au cours de la recherche. Dans mon cas, le but est d'obtenir un format de quiz efficient et donnant des résultats optimaux, avec un minimum de temps d'investissement de l'enseignant. J'ai donc adopté une posture de type chercheur acteur, où j'étais partie prenante des changements (Karsenti & Savoie-Zajc, 2004).

3.1.2 Nature du corpus

3.1.2.1 Population

Le corpus est identique à celui de la première partie de mon étude, à savoir une classe d'allemand de 9e Harmos composée de 20 élèves (11 garçons et 9 filles).

3.1.2.2 Procédure et protocole de l'expérience

Suivant la structure proposée par Dolbec & Clément (2004), la recherche a revêtu la forme d'un processus cyclique d'action et de réflexion. Une phase de mise en application d'une action en vue de produire un changement a été suivie d'une prise de recul pour évaluer l'efficacité du dispositif et préparer l'action suivante.

Dans les faits, les impressions des élèves ont été récoltées chaque semaine, juste après la passation du quiz. Diverses questions leur ont été posées, relatives à l'ergonomie, la simplicité et l'accessibilité du dispositif, ainsi que sur leur ressenti. En parallèle, j'ai fait un bilan personnel relatif au quiz effectué, en prenant en compte l'aisance et le temps nécessaire pour la création du test, la facilité de déploiement du dispositif, les réactions des élèves pendant le quiz, ainsi que leurs résultats.

Quant aux questions adressées aux élèves, je me suis inspiré pour certaines d'entre elles du TER de Bieri (2016, p. 12), notamment concernant la clarté des questions, la préférence entre format papier ou numérique, ou encore du délai du feedback proposé.

Grâce à ces données, l'idée était que le dispositif évolue de semaine en semaine, par exemple en procédant à des modifications d'interface (police de caractère, couleurs, présence ou non d'images), en variant l'ordre des questions, la manière dont le feedback était communiqué, ou encore le contenu de ce dernier.

3.1.2.3 Problématique de l'accès à Internet

Les quiz étant hébergés sur le site ProProfs.com, un accès Internet était nécessaire pour pouvoir y accéder. Seule une minorité d'élèves bénéficiait d'un forfait mobile comprenant un accès 4G illimité, il a donc fallu trouver une solution pour offrir un accès Wifi aux élèves ne possédant pas une telle prestation. Au sein de mon école, il n'y a pas d'antenne Wifi à disposition des élèves ; un tel dispositif ne sera installé que pour la rentrée 2017/2018. Il aurait été possible de brancher un router wifi à la prise réseau présente dans ma salle de classe, mais dans un souci de simplicité et toujours dans l'optique de permettre à n'importe quel enseignant de mettre sur pied un dispositif similaire à l'avenir, j'ai opté pour la création d'un point d'accès Internet à l'aide de mon téléphone portable. Ce procédé consiste à partager la connexion 4G d'un téléphone en permettant à d'autres appareils de se connecter via Wifi ; les appareils connectés ont ainsi accès à Internet via le réseau mobile du téléphone faisant office de passerelle. Le système d'exploitation Android limitant l'accès à 10 appareils, il a fallu utiliser un deuxième point d'accès Internet pour garantir aux 20 élèves de ma classe l'accès à Internet. Une élève au bénéfice d'un accès 4G illimité ayant spontanément proposé de partager sa connexion lors d'un test en amont de l'étude, j'ai décidé d'opter pour cette solution.

Cette dernière a pour avantages son prix nul (pour autant que l'enseignant soit au bénéfice d'un abonnement 4G suffisant) et sa rapidité de mise en place (sans matériel additionnel). De plus, les élèves sont habitués à partager leur connexion entre eux ; ils n'ont donc ainsi eu aucun problème à se connecter au dispositif.

Cependant, le partage d'une connexion privée a quelques désavantages par rapport à une connexion sécurisée d'établissement (telle que celle fournie par le RPN, notamment). Un blocage de certains sites via un serveur proxy n'est pas possible ; les élèves sont donc susceptibles d'accéder à des sites interdits. Il leur est également possible d'aller chercher des informations en ligne sur Google ou sur un dictionnaire en ligne, en vue de compléter les quiz. Je ne peux pas garantir que les élèves ne se soient

pas rendu sur des sites autres que celui dédié aux quiz pendant leur passation, mais le fait de circuler dans la classe, de demander aux élèves ayant terminé de ranger leurs téléphones, et de couper le partage de connexion dès la fin du quiz m'ont semblé être des précautions suffisantes.

3.1.2.4 Accès aux quiz

Une fois l'accès Internet résolu, il m'a également fallu trouver un moyen de permettre aux élèves d'accéder aux quiz aisément, sans devoir en saisir l'adresse URL complète. En effet, celle-ci revêtait la forme ci-dessous :

<http://www.proprofs.com/quiz-school/story.php?title=review-quiz-2>

Si les élèves avaient dû saisir cette adresse, cela aurait pris du temps et le risque d'erreur aurait été important. Afin d'éviter cela, j'ai envoyé le lien sur le groupe WhatsApp⁹ de la classe et les élèves n'ont eu qu'à cliquer sur ce lien pour accéder au portail de quiz. Pour les élèves n'ayant pas WhatsApp et n'étant de ce fait pas dans le groupe de la classe, j'ai utilisé le réducteur d'URL bitly¹⁰ afin d'obtenir des adresses plus courtes, du type :

<http://bit.ly/2IPvcFp>

J'ai ensuite affiché cette adresse à l'écran, via le beamer. Bien qu'ayant bien fonctionné, ce type d'URL a pour défaut d'utiliser indifféremment des caractères en majuscule et minuscule, ce qui peut entraîner des erreurs de frappe. Certains élèves ont rencontré ce problème et je suis passé vers eux afin d'écrire l'URL correcte. Après plusieurs semaines, ce problème a cependant disparu, les élèves ayant pris l'habitude de veiller à respecter la casse des caractères.

3.1.2.5 Disponibilité des téléphones portables des élèves

Un autre point à prendre en compte lors de la mise en place d'un dispositif de quiz dans une classe du secondaire 1 est lié à la disponibilité des téléphones portables des élèves. Malgré l'envoi d'une circulaire explicative en amont aux parents leur demandant d'autoriser leurs enfants à prendre leur téléphone en classe d'allemand pendant la durée de l'étude, je n'ai pas pu compter sur la totalité des smartphones en classe.

⁹ Système de messagerie instantanée par Internet sur mobile utilisé par 1 milliard de personnes.

¹⁰ Accessible en ligne à l'URL <https://bitly.com>

En effet, entre ceux qui avaient oublié le leur chez eux ou ceux dont la batterie était insuffisamment chargée pour naviguer sur Internet, il n'y a pas eu une fois où l'ensemble des élèves de la classe était en possession d'un téléphone portable fonctionnel. Pour palier ce problème, j'ai amené en classe une tablette, un téléphone portable additionnel, et ai mis à disposition des élèves l'un des téléphones utilisés comme point d'accès Internet en cas de besoin. Ainsi, je bénéficiais de 3 périphériques additionnels pour mener à bien mon dispositif. Dans le cas où je n'aurais pas eu en ma possession ce matériel, l'école possède des iPod Touch permettant l'accès à Internet par Wifi, que j'aurais pu réserver et utiliser si besoin.

3.1.3 Méthode de collecte des données

Vu le nombre restreint de participants à l'étude, j'ai décidé de collecter les données via la méthode du focus group. Kitzinger et al. (2004) définissent les focus groups comme des discussions de groupe ouvertes, organisées dans le but de cerner un sujet ou une série de questions pertinents pour une recherche. J'ai choisi cette méthode de récolte de données pour des questions de temps, car des méthodes telles que les questionnaires et entretiens auraient été longs à mettre en place et auraient pris plus de temps à analyser. J'ai ainsi pu récolter rapidement le retour des élèves et modifier en conséquence mon dispositif de semaine en semaine.

La récolte des données a été effectuée via la méthode du focus group, à raison d'une fois par semaine, juste après la passation du quiz hebdomadaire. La durée prévue à cet effet était d'environ 5 minutes, juste avant la fin de la leçon. Le protocole mis en place a été constant durant toute la durée de l'étude et comportait les étapes suivantes :

1. Vers la fin du test, affichage des questions au beamer, afin que les élèves qui avaient déjà terminé le quiz puissent y réfléchir et prendre des notes.
2. Retour sur les points problématiques ou pistes d'amélioration évoquées lors du focus group précédent afin de voir si les modifications apportées au dispositif avaient porté leurs fruits.
3. Discussion en plenum sur les questions affichées. Plusieurs avis ont été récoltés pour chaque point. Les élèves interagissaient s'ils le souhaitaient, mais j'ai cependant profité d'interroger quelques élèves au hasard sur les différents points.
4. Prise de note des éléments positifs, des éléments à améliorer, ainsi que des autres remarques qui auraient pu survenir.

3.2 Analyse

Au contraire des données récoltées pour la première partie de mon étude, qui nécessitaient diverses opérations pour pouvoir être analysées, les informations collectées en focus group ont été pour moi directement utilisables pour améliorer mon dispositif de semaine en semaine ; comme elles ont été collectées oralement et non à l'aide d'un questionnaire, elles ne nécessitaient pas de traitement complémentaire pour être exploitables. De plus, la discussion des diverses questions en plenum a permis de faire émerger une impression globale de la classe concernant chaque point.

3.2.1 Résultats et changements apportés

Après chaque retour des élèves, les informations récoltées ont été placées dans un document que je reprenais ensuite chaque semaine lors de la création du quiz, afin de tenir compte des remarques formulées. La section ci-après regroupe les résultats obtenus au fil des semaines, tant au niveau des observations en classe que du retour en focus group, ainsi que les améliorations apportées au dispositif chaque semaine :

1ère semaine

Quiz 1	
Nombre de soumissions	18
Nombre de questions	16
Score moyen	33%
Temps de passation moyen	07:45

Tableau 4: résumé du quiz 1

Déroulement du quiz 1

L'absence imprévue de l'élève censée partager sa connexion 4G avec 10 camarades a entraîné un problème de connexion Internet et seuls 10 élèves ont pu effectuer le quiz. Par chance, j'ai pu profiter du fait que la leçon suivante était également de l'allemand, mais en demi-groupe, pour faire passer le quiz de manière individuelle aux autres élèves, sur le poste du maître cette fois et non pas sur leurs smartphones. Pendant la passation du quiz, de nombreuses demandes de clarification de données ou de questions ont été formulées par les élèves.

Questions posées aux élèves

- Interface de l'outil : conviviale / austère ?
- Difficultés techniques rencontrées ? Clavier ?
- Qualité de l'explication donnée en cas de mauvaise réponse: bonne / à améliorer ?
- Images : plus facile avec / ne change rien ?
- Images : suffisamment claires ?

- Ordre des questions : ordrées / mélangées ?
- Dernière page : bien d'avoir ses résultats directement?
- Utilité d'un tel quiz ?

Retour du focus group (9 élèves durant la leçon suivant le quiz, en demi-groupe)

- Interface : celle proposée par l'outil ProProfs est plaisante et est préférable à des fiches papiers. Cette remarque était partagée par la totalité des élèves, sauf un qui préfère les fiches papier. L'élève en question, habitué à de bonnes notes (moyenne annuelle de 5.36) a réalisé un score moyen pour lui de 37.5% à ce premier quiz. Par la suite, il s'est amélioré à chaque quiz, jusqu'à obtenir une moyenne de 90.38% à l'évaluation finale. Cet élève ayant manqué le quiz 4, il n'a pas été pris en compte dans les résultats de mon étude.
- Difficultés techniques : certains ont trouvé que les trémas étaient difficiles à réaliser, mais à l'unanimité ils préfèrent les écrire via un clavier de téléphone portable plutôt que sur un clavier d'ordinateur.
- Feedback donné : la qualité est bonne et permet de comprendre ses erreurs. A la question de savoir s'il apporte quelque chose, les élèves ont relevé à l'unanimité son importance. Le fait de recevoir un feedback immédiat dès la fin du test est également une très bonne chose. Dans l'idéal, au lieu des appréciations, ils souhaiteraient avoir directement l'équivalence en notes.
- Images : elles améliorent la perception positive du quiz et permettent de bien situer le contexte. Les images choisies étaient selon eux suffisamment claires, sauf celles montrant des directions.
- Ordre des questions : les élèves préfèrent les avoir dans le désordre, cela donne de la diversité.
- Utilité d'un quiz : à l'exception d'un élève, tous ont eu l'impression que cela leur était utile, car cela leur permettait de voir où ils en étaient dans leur apprentissage.
- Affichage des questions : bien d'en avoir une seule par page, pour des raisons d'ergonomie. Par contre, les élèves auraient préféré pouvoir revenir en arrière si nécessaire (ce qui n'était pas possible dans le quiz proposé) ou d'avoir une révision finale avant de soumettre leur réponse.

Réflexion personnelle suite au quiz 1

L'outil ProProfs est aisé à prendre en main et permet facilement de créer des questions de type QCM et choix ouvert, ceci même pour un non spécialiste. En ce sens, les observations après le premier quiz m'ont conforté dans l'idée que l'outil sélectionné répondait autant à mes besoins qu'aux attentes des élèves. J'ai donc pu continuer dans cette voie.

Le problème d'accès Internet rencontré a dû être solutionné pour la semaine d'après. Le fait que chaque élève visualisait les questions dans un ordre différent leur évitait d'être tenté de tricher sur le camarade d'à côté. Au vu des nombreuses questions des élèves pendant le quiz, un travail sur la formulation des consignes a été réalisé. Au fil des semaines, les élèves se sont cependant habitués au mode de questionnement de type quiz et les questions ont diminué. Concernant le temps alloué, certains élèves ont semblé prendre tout leur temps pour compléter le quiz, alors que d'autres avaient terminé après quelques minutes seulement.

Choix des modifications à apporter suite au quiz 1

- Conversion du feedback sous forme de pourcentage et d'appréciation au format de note, en fonction du pourcentage de bonnes réponses données (100-90% => 6, 80-89% => 5.5, 70-79% => 5, 60-69% => 4.5, 50-59% => 4, 40-49% => 3.5, 30-39% => 3, 20-29% => 2.5, 10-19% => 2, 0-9% => 1.5).
- Ajout d'un minuteur limitant la passation du quiz à 10 minutes (but : éviter une baisse de concentration et une dissipation des élèves, ne pas perdre trop de temps non plus).
- Présentation des questions : ajout d'un titre indiquant de quoi il s'agit (vocabulaire, prépositions, directions, etc.) et ajout de données en français, afin de s'assurer que chacun comprenne.
- Accès à Internet : au vu du problème rencontré suite à l'absence de l'élève censée partager sa connexion, j'ai décidé d'utiliser le téléphone portable de mon épouse, que je remercie au passage, pour créer ce second accès. Bien que répondant aux besoins de mon étude, le dispositif mis en place n'est à mon sens pas optimal car il ne permet pas de filtrer l'accès des élèves à certains sites, à la manière d'un serveur proxy.

2e semaine

Quiz 2	
Nombre de soumissions	18
Nombre de questions	15
Score moyen	32%
Temps de passation moyen	06:46

Tableau 5: résumé du quiz 2

Déroulement du quiz 2

Les problèmes de connexion rencontrés la première semaine ont été résolus par l'apport d'un deuxième smartphone partageant sa connexion, permettant ainsi à 20 élèves de se connecter. 2 élèves étaient absents ce jour-là, 18 élèves ont passé le quiz.

Questions posées aux élèves

- Difficultés techniques rencontrées ? Clavier ?
- Qualité de l'explication donnée en cas de mauvaise réponse : bonne / à améliorer ?
- Images : suffisamment claires ?
- Dernière page : bien d'avoir ses résultats directement ? Note ?
- Ajout du timer : positif / négatif ?
- Autres commentaires ?

Retour du focus group (18 élèves, en classe complète)

- Le clavier du téléphone n'a plus été un obstacle pour la grande majorité des élèves. Ils ont su comment obtenir un tréma et ont pris conscience d'activer ou de désactiver les majuscules automatiques.
- Le texte donné comme feedback était suffisant et adapté.
- Les images utilisées étaient claires et pertinentes. Attention cependant à leur taille, qui peut rendre certaines informations difficiles à lire.
- Feedback immédiat sous forme de note : plus intéressant que des lettres ou des appréciations. Par contre, cette note n'est pas assez mise en évidence sur la page finale de résultats.
- Durée limitée à 10 minutes : cela engendre un stress chez les élèves. Certains ont bâclé les dernières questions car ils arrivaient au bout du temps alloué. 1 élève n'a pas pu terminer et s'est vu assigner un résultat de 0%. Pour cette raison, ces résultats n'ont pas été pris en compte pour cette étude.
- Présentation des questions : plus claire que la semaine passée, les élèves savaient naturellement ce qui était attendu, notamment pour les questions courtes.

Réflexion personnelle suite au quiz 2

Le protocole d'accès à Internet a fonctionné et s'est avéré aisé à mettre en place. L'ajout d'un timer a provoqué chez certains élèves une plus grande application, mais chez d'autres une panique et une frustration lorsqu'ils n'ont pas pu terminer le quiz. L'interface épurée proposée (texte noir sur fond blanc) pourrait susciter de la lassitude chez les élèves si aucune variation n'était apportée les prochaines semaines.

Choix des modifications à apporter suite au quiz 2

- Retirer le timer : celui-ci engendre du stress et fausse les résultats si un élève n'a pas le temps de terminer.
- Suite du travail sur la présentation du contenu, les données et la formulation des questions.

- Modification du thème de l'interface : fond bleu et police blanche, afin de voir si le ressenti des élèves est positif ou négatif.

3e semaine

Quiz 3	
Nombre de soumissions	18
Nombre de questions	13
Score moyen	33%
Temps de passation moyen	11:31

Tableau 6: résumé du quiz 3

Déroulement du quiz 3

17 élèves étaient présents ce jour-là. 3 élèves absents n'ont pas passé le quiz. J'ai alloué 20 minutes en fin de leçon pour passer le test mais n'ai plus eu de temps pour la partie focus group en plenum, je me suis donc rabattu sur le demi-groupe qui suivait.

Les élèves ont posé moins de questions pendant le déroulement du quiz ; ceux qui ont continué à poser des questions étaient les élèves les plus faibles de la classe, qui n'étaient souvent pas capables de comprendre la consigne ou la phrase à compléter.

Questions posées aux élèves

- 4 choix pour le QCM : trop facile / correct / trop dur ?
- A quel type de question est-il plus facile de répondre ?
- Quel type de question sollicite plus vos connaissances ?
- Titres et consignes : clairs ?
- Couleur de l'interface : blanc ou mieux en couleur ?
- Police utilisée : standard ou fantaisiste ?
- Contenu proposé : bon résumé de ce qui a été fait ?

Retour du focus group (10 élèves durant la leçon suivant le quiz, en demi-groupe)

- Forme du QCM : 4 réponses possibles était adapté et a été apprécié des élèves
- Le QCM était plus facile à répondre pour eux
- Ils étaient conscients que les réponses courtes nécessitent plus de travail de récupération.
- Les titres et les consignes étaient plus clairs et n'ont plus nécessité d'explication complémentaire chez la majorité des élèves.
- Le changement de thème a bien été accueilli. Le fond bleu était plaisant et faisait mieux ressortir l'écriture et les images. Sur le choix de la police, les élèves étaient plus mitigés :

certains auraient aimé des polices plus fantaisistes, d'autres ne souhaitaient pas changer.

- Images : certaines étaient très petites et pas toujours pertinentes. Selon le smartphone utilisé, certains n'avaient pas la possibilité de zoomer pour agrandir l'image.
- Le contenu proposé était un bon résumé de ce qui a été fait dans la semaine et leur a permis de mesurer leur apprentissage.
- Le fait de retirer le timer a été apprécié.
- Interface : les élèves auraient souhaité pouvoir retourner en arrière, comme cela aurait été le cas au format papier.

Réflexion personnelle suite au quiz 3

Au vu du faible pourcentage de réussite moyen aux quiz effectués jusqu'à présent, je me suis demandé s'il n'était pas judicieux de faire varier le mode de feedback proposé, en n'attendant plus la fin du quiz mais en l'affichant après chaque réponse. En ajoutant cette modification, il n'était pas possible d'offrir la possibilité de naviguer entre les questions, comme les élèves le souhaitaient, car cette fonctionnalité aurait permis de revenir sur des questions dont la réponse et le feedback avaient déjà été affichées.

Je me suis aussi demandé si de meilleurs résultats seraient obtenus en variant le nombre de questions proposées.

Choix des modifications à apporter suite au quiz 3

- Aggrandir la taille des images proposées et s'assurer que leur qualité soit suffisamment pertinente.
- Ajouter le feedback immédiat (affiché après chaque question).
- Varier le nombre de questions pour voir si les résultats s'améliorent (diminution à 9 questions).

4e semaine

Quiz 4	
Nombre de soumissions	17
Nombre de questions	9
Score moyen	42%
Temps de passation moyen	06:40

Tableau 7: résumé du quiz 4

Déroulement du quiz 4

17 élèves étaient présents ce jour-là. 3 élèves absents n'ont pas passé le quiz.

Les élèves ont eu 10 minutes en milieu de leçon pour réaliser le quiz. Concernant les questions pendant le quiz, même remarque que la semaine précédente, ceux qui en ont posé étaient les élèves qui ne lisent pas les consignes ou qui ne n'étaient pas (encore) capables de bien comprendre ce qu'est un déterminant, une préposition, etc.

Questions posées aux élèves

- Mode de feedback : mieux direct ou différé ?
- Nombre de questions : mieux lorsqu'il y en a moins que 10 ?
- Motivation
- est-ce qu'un quiz est motivant ?
- par rapport au papier, en quoi cela change-t-il de connaître son résultat directement ?
- est-ce que vous comparez vos scores entre vous ?
- Autres quiz dans d'autres branches ?
- Si non, dans lesquelles en aimeriez-vous ?

Retour du focus group

Le feedback immédiat a été apprécié par la totalité des élèves. Il a diminué leur anxiété et a permis d'éviter d'autres erreurs similaires par la suite, notamment celles liées au type de contenu attendu pour les réponses courtes. Le nombre de questions en diminution a également été salué, cela a permis d'avoir un quiz plus court et moins exigeant pour les élèves. Au niveau de la motivation, les élèves ont trouvé que passer un quiz était motivant, notamment parce qu'il permet de sortir et d'utiliser leur téléphone en classe, ce qui n'est pas courant. Les quiz ont pour avantage de livrer un résultat immédiat, au contraire de tests papier, ce que les élèves ont beaucoup apprécié. En fin de test et parfois également après à la pause, ils ont échangé leurs scores et en parlaient entre eux. Aucun autre enseignant de l'école ne pratique les quiz, mais les élèves verraient bien l'émergence d'une telle pratique en français, notamment pour les aider avec les questions de conjugaison.

Réflexion personnelle suite au quiz 4

Au vu du plébiscite des élèves pour l'utilisation du smartphone en classe, je souhaitais également voir si un quiz passé sur un ordinateur était apprécié.

Choix des modifications à apporter suite au quiz 4

Conservation du feedback immédiat après chaque question, à la demande des élèves. Afin de voir si un changement d'interface a une incidence sur la perception des élèves, j'ai décidé de faire passer le test de révision de la semaine 5 en salle d'informatique, sur des ordinateurs standards.

5e semaine

Quiz de révision	
Nombre de soumissions	13
Nombre de questions	53
Score moyen	46%
Temps de passation moyen	25:12

Tableau 8: résumé du quiz 5

Déroulement du quiz de révision

Le quiz s'est déroulé en salle d'informatique, sur une période. En raison d'actions que je n'ai pas été en mesure d'identifier sur le moment, les pages de certains élèves se sont rechargées pendant le quiz, les obligeant à recommencer depuis le début. Cela peut être dû à une mauvaise manipulation au clavier de leur part, à un bug du navigateur utilisé (Internet Explorer), ou de la version ordinateur de l'outil ProProfs. A cause de ces soucis, plusieurs élèves n'ont pas pu terminer le quiz, mais avaient déjà répondu à plus de 80% des questions à la fin de la leçon. En partant du principe qu'ils ont été exposé à d'autres questions avant le rechargement de leur page, j'ai décidé de prendre en compte leurs résultats pour mon étude.

Question posée aux élèves

- Au niveau de l'interface, est-il préférable d'utiliser l'outil via un smartphone ou un ordinateur ?

Retour du focus group

Les réponses ont fortement divergé. Une majorité des élèves penchait néanmoins pour la version smartphone. Plusieurs élèves, à l'aise avec l'outil informatique, ont cependant manifesté une préférence pour la version ordinateur, mettant en avant les avantages de la taille de l'écran, du clavier plus facile à utiliser et des images plus lisibles.

3.2.2 Evolution des questions posées

Au fur et à mesure de l'avancement de l'expérimentation en classe, de mes besoins en termes de retour des élèves, des retours précédents ainsi que de mon ressenti sur le déroulement des quiz, le type de questions posées aux élèves a évolué, selon le schéma ci-après :

Première semaine :

- questions sur l'outil (interface de l'outil austère ou conviviale / difficultés techniques rencontrées : connexion, accès au quiz, utilisation du quiz, utilisation du clavier / avantages

d'avoir la page de résumé comportant les résultats directement en fin de quiz).

- questions sur le contenu : présence des images positive ou non, pertinence des images proposées, ressenti sur l'ordre des questions aléatoire, qualité de l'explication donnée en cas de mauvaise réponse.
- questions sur la démarche : pertinence et utilité d'un tel quiz.

Deuxième semaine :

- retour sur d'éventuelles questions liées aux difficultés techniques.
- questions sur l'outil (ordre des questions, problèmes de clavier, notes en dernière page)
- questions sur le contenu (pertinence des images proposées)
- questions générales : autres commentaires.

Troisième semaine :

- questions sur le type de questions proposées : nombre d'éléments à choix pour le QCM.
- questions sur la présentation du contenu : titres, consignes (indications préalables à l'exercice), données, couleurs de l'interface, polices.
- questions sur le contenu : résume bien ce qui a été vu dans la semaine ?
- questions sur l'expérience : quel type de question est plus facile à répondre ? Quel type de question sollicite plus vos connaissances/demande un effort mental ?

Quatrième semaine :

- questions sur le mode de feedback (immédiat vs différé)
- questions sur le nombre de questions (mieux <10 ou 15)
- questions sur la motivation :
- est-ce motivant d'avoir un quiz ?
- est-ce que vous comparez vos scores avec ceux des autres ?
- qu'est-ce que cela change d'avoir son score immédiatement ?
- est-ce que des enseignants d'autres branches en font également ?
- aimeriez-vous en avoir dans d'autres branches ? Si oui, lesquelles ?

Cinquième semaine (review test) :

- question sur le périphérique utilisé pour la passation du quiz (mobile / ordinateur)

Comme vous pouvez le constater, les questions initiales quant à la pertinence du choix de l'outil et l'intérêt de la démarche en soi ont rapidement disparu au profit d'interrogations liées aux contenus et à la présentation de ceux-ci, ainsi qu'au ressenti des élèves.

3.2.3 Discussion

La recherche-action sous forme de processus cyclique d'action et de réflexion s'est avérée être un bon choix, notamment au vu de la récolte directe et aisée des données via la méthode du focus group, et l'évolution rapide du dispositif au fil des semaines. Néanmoins, le défaut des focus groups est que seulement certains élèves donnent leur avis; on ne récolte ainsi pas l'avis de tous les élèves. De plus, la période de 5 semaines sur laquelle s'est déroulée l'étude est très courte pour permettre une évolution profonde du dispositif. Il serait souhaitable de le tester sur au minimum un autre chapitre (5 autres semaines) pour véritablement proposer des changements en profondeur et mesurer leurs effets.

La plateforme ProProfs utilisée pour l'étude a tenu toutes ses promesses en termes de facilité d'utilisation, robustesse et utilisation des données après passation des tests. Son défaut reste le fait qu'elle soit payante, ce qui pourrait décourager des utilisateurs éventuels au profit d'autres outils tels que Google Forms.

L'idée de base du dispositif était d'éviter un déplacement en salle d'informatique ou l'acquisition de matériel dédié pour la passation des quiz en utilisant les téléphones portables des élèves. Bien qu'ayant créé quelques soucis lors de la mise en place effective en classe, cette manière de procéder a bien fonctionné et a été très appréciée des élèves.

Le fait que l'interface de l'outil s'adapte aux différents périphériques (smartphones, tablettes, ordinateurs) a permis de tester le dispositif dans différentes conditions et, bien que la majorité des élèves ait préféré compléter les quiz sur leurs smartphones, il est fort probable que leur motivation était biaisée par le fait de pouvoir utiliser leur téléphone en classe, ce qui d'ordinaire ne leur est pas autorisé. A l'avenir, je pense qu'il serait intéressant de ne pas les obliger à utiliser un support défini, mais de leur laisser le choix, selon leurs préférences et affinités avec l'un ou l'autre périphérique.

4. Pistes d'amélioration du dispositif

Au vu des résultats obtenus et des remarques formulées tout au long de ce travail, il apparaît que le dispositif mis en place est perfectible sur divers points.

D'abord, le fait qu'il n'ait pas été bénéfique pour tous les élèves est problématique et mérite d'être investigué. Une raison possible de la baisse de performance chez certains élèves lors de l'évaluation finale pourrait résider dans le contenu et le format de ladite évaluation. En effet, au contraire des épreuves d'allemand administrées depuis le début de l'année, qui portaient sur des aspects communicationnels (compréhension orale et écrite, expression orale et écrite, le tout dans des contextes d'utilisation précis), l'évaluation sommative proposée dans le cadre de cette étude portait principalement sur des aspects grammaticaux et lexicaux. Au niveau de la présentation des questions également, le format sous forme de choix ouvert, bien qu'exercé à de multiples reprises dans les quiz, était inhabituel pour une épreuve. Ces deux aspects sont susceptibles d'avoir déstabilisé les élèves, en particulier ceux rencontrant déjà des difficultés en allemand. Une piste d'amélioration consisterait à continuer de pratiquer des tests de rappel chaque semaine pendant une séquence d'allemand, mais en laissant l'évaluation finale dans un format plus traditionnel, afin de voir si les résultats sont meilleurs. Normalement, les questions des quiz n'ayant pas d'impact direct dans l'évaluation mais agissant simplement comme des aides à la mémorisation en cours de séquence, on s'attendrait à ce que les résultats des élèves soient positivement influencés ou neutres, mais pas péjorés.

Un second point concerne un manquement des outils proposés : autant dans Google Forms que dans ProProfs, il n'y a actuellement pas de fonctionnalité permettant aux élèves de reprendre des items qui n'auraient pas bien fonctionné pour les retravailler par la suite, comme c'est le cas dans Quizlet¹¹ avec la fonction flashcards. Il serait intéressant d'avoir un outil qui permette de classer les items selon qu'ils sont « acquis » ou « à revoir », ou de pouvoir revenir sur des contenus de manière périodique. ProProfs comporte de nombreux produits additionnels à l'outil « Quiz Maker » utilisé dans cette étude, notamment un outil de création de flashcards. Celui-ci est cependant également payant – contrairement à Quizlet – et ne permet pas de réutiliser des items utilisés dans des quiz ; seules des listes de correspondance mot-mot ou mot-image peuvent être utilisées. Concernant Google Forms, aucune fonctionnalité permettant de présenter à nouveau du contenu n'ayant pas bien fonctionné à un apprenant n'est présente, mais le fait qu'une API¹² existe pour ce produit pourrait permettre à un

¹¹ Outil d'aide à l'apprentissage du vocabulaire, accessible à l'adresse www.quizlet.com

¹² Acronyme de « Applications Programming Interface » : il s'agit d'une interface de programmation qui permet de se connecter à une application pour échanger des données.

développeur d'intégrer ce type de fonctionnalité via un *add-on* à l'avenir. Finalement, une piste de solution pourrait bien provenir de Quizlet qui a annoncé début mars sur son blog la sortie de Quizlet Learn, un outil d'apprentissage qui permet l'élaboration d'un plan adapté en fonction d'une échéance (Glotzbach, 2017). L'ouvrage *Make it Stick* (Brown, 2014) est explicitement cité comme source d'inspiration pour cet outil, qui repose sur l'assistant d'apprentissage de Quizlet, une plate-forme qui analyse des millions de séances de révision anonymes et qui met en œuvre les résultats à l'aide de notions tirées du domaine des sciences cognitives. Cet outil, bien que limité à l'apprentissage de listes de mots, notamment de vocabulaire pour les langues, semble donc comporter un énorme potentiel et il sera intéressant de garder un œil attentif sur son évolution ces prochains mois.

L'aspect de la problématique de l'accès à Internet en classe est également à améliorer. La solution de partage de connexion mise en œuvre, bien qu'aisément déployable, n'est pas totalement satisfaisante car elle ne permet pas de filtrer l'accès à certains sites. Il serait souhaitable de bénéficier d'un accès Wifi pour les élèves au travers d'une plateforme sécurisée empêchant l'accès à certains sites, à la manière d'un serveur proxy. Dans l'idéal, il faudrait que l'enseignant puisse indiquer les sites autorisés, de manière à ne permettre l'accès qu'aux quiz. Des pistes de solution dans ce sens semblent émerger sur le canton de Neuchâtel, avec la mise à disposition de bornes Wifi dédiées aux élèves sur demande des directions. Dans mon école, ce sera le cas pour la rentrée 2017-2018 et il sera intéressant de voir les possibilités de paramétrage de ces bornes.

Le dernier point problématique du dispositif de tests mis en place a été la présence des élèves en classe. Dans le modèle choisi, la présence des élèves en classe au moment choisi pour la passation d'un quiz était requise. Dans les faits, de nombreux élèves (40% de l'effectif) ont manqué un ou plusieurs quiz pour divers motifs (maladie, rendez-vous médical, congés accordés par la direction, etc.). Sachant cela, l'utilisation des questions provenant des quiz intermédiaires pour l'évaluation finale n'est pas adaptée, car les élèves non présents n'ont pas accès aux questions des quiz en-dehors de leur passation, à l'exception du quiz de révision, qui comporte l'ensemble des questions. Afin de palier ce problème, il serait judicieux de permettre aux élèves de se tester en-dehors des cours via la plateforme de quiz ou, comme proposé plus haut, ne pas réutiliser les questions des quiz dans l'évaluation finale mais utiliser celles-ci dans une perspective de mémorisation seulement.

Conclusion

Le dispositif mis en place dans une classe d'allemand de 9^e Hamos a permis de démontrer l'existence d'un *testing effect* substantiel dans une classe de langue seconde du secondaire 1, pour des questions de type QCM et choix ouvert, par rapport à la matière non préalablement testée. Ces résultats permettent de confirmer l'hypothèse de départ qui était que les conclusions de l'étude de Roediger III et al. (2011), obtenues dans un cours de sciences sociales au sein d'une *middle school* aux Etats-Unis, se vérifiaient dans un cours de langue du secondaire 1. Le dispositif n'a cependant pas été bénéfique pour la totalité des élèves. En effet, un tiers des participants a enregistré une note inférieure à sa moyenne annuelle lors de l'évaluation finale, avec une baisse de performance de 24.13% en moyenne. En comparaison, les élèves qui ont augmenté leurs performances grâce au dispositif l'ont fait en moyenne de 3.13%. Ces derniers résultats sont cependant à prendre avec des pincettes car l'application du test de Student sur les données a démontré qu'ils ne sont pas statistiquement significatifs. En ce sens, il serait judicieux de tester le dispositif à plus large échelle afin d'obtenir des résultats généralisables. La variation de mode de feedback au cours du quatrième quiz a été favorablement accueillie par les élèves mais son impact n'a pas pu être mesuré au cours de la présente étude. Il serait intéressant de mener une étude sur cet aspect uniquement, afin de voir l'impact de timings différenciés du feedback.

La seconde partie de l'étude visait à définir un dispositif permettant d'intégrer des tests à la pratique existante d'un enseignant, sans nécessiter un investissement de temps massif de ce dernier. En ce sens, l'utilisation de l'outil ProProfs à l'aide des smartphones personnels des élèves représente une solution très satisfaisante et bon marché. L'interface offerte, de type *responsive*, permet une utilisation sur tous types de périphériques (tablettes, ordinateurs, mobiles), offrant ainsi une grande flexibilité à l'enseignant quant au placement des quiz (en classe, en salle d'informatique, ou à la maison). La fréquence hebdomadaire choisie pour les quiz est également judicieuse, autant du point de vue de l'enseignant que des élèves. En effet, la quantité de manière testée n'est ainsi pas trop conséquente et les élèves peuvent ainsi, au travers d'un nombre de questions restreint, monitorer leurs apprentissages et leur compréhension de la matière. La forme de feedback choisie, différée en fin de test, n'était apparemment pas idéale. Selon les retours des élèves lors des focus groups, un feedback immédiat après chaque question aurait été préférable ; il s'agit là d'une piste intéressante pour l'amélioration du dispositif. Cette seule variation de l'étude pourrait faire l'objet d'un travail de recherche complémentaire intéressant. Dans tous les cas, ce sujet mérite d'être investigué. L'objectif initial étant d'offrir à un enseignant lambda un dispositif permettant la mise en place de tests de rappels à un enseignement existant, il serait intéressant de proposer le dispositif à des collègues, de langues secondes ou d'autres branches, et de leur demander de l'expérimenter.

Dans notre vie d'enseignant, nous sommes en permanence dans une posture de recherche, afin de promulguer un enseignement efficace pour la majorité des élèves. Dans le canton de Neuchâtel, avec la réforme du cycle 3 de l'école obligatoire qui implique un grand nombre de cours donnés en classes hétérogènes, un concept est sur toutes les lèvres : la différenciation pédagogique. Selon Laurent (2001), la pratique de la différenciation pédagogique consiste à organiser la classe de manière à permettre à chaque élève d'apprendre dans les conditions qui lui conviennent le mieux. Il ne s'agit pas de différencier les objectifs, mais de permettre à tous les élèves d'atteindre les mêmes objectifs par des voies différentes. La difficulté pour les formateurs consiste à ne pas simplement répéter d'une autre manière, mais à « varier le plus possible leurs actions, pour que chacun puisse rencontrer, à un moment ou l'autre de son cursus, des situations dans lesquelles il puisse réussir » (Fournier, 1996). A terme, mettre à disposition des tests différents selon les besoins des élèves serait, à mon sens, un moyen efficace de différencier.

Ce travail de recherche me permettra de mettre en place périodiquement dans mes classes de petites recherches afin d'améliorer la qualité de mon enseignement. En ce sens, la connaissance d'éléments statistiques me sera d'une grande utilité pour analyser les données récoltées. Le fait de se baser sur un modèle scientifique éprouvé pour tester différentes hypothèses et variables dans mes classes renforcera la crédibilité des résultats et permettra ainsi un meilleur partage auprès d'autres enseignants ou cercles scolaires.

A terme, je souhaiterais mettre à disposition des élèves un outil informatisé ou une liste de liens en relation avec le thème du chapitre afin de les inciter à se tester eux-mêmes en-dehors du cours. L'idée serait d'augmenter le pourcentage de la matière testée et le nombre de tests passés, ainsi que de permettre aux élèves de se tester et, de ce fait, prendre conscience de leurs forces et faiblesses, au travers d'un outil commun.

Références bibliographiques

- Bieri, J. (2016). *Dans quelle mesure des quiz distribués aux élèves sous format informatique permettent-ils une meilleure rétention de l'information dans une classe de 10H en sciences de la nature?* (Travail écrit de recherche). HEP-BEJUNE.
- Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, 8(3), 323–343. [https://doi.org/10.1016/S0022-5371\(69\)80124-6](https://doi.org/10.1016/S0022-5371(69)80124-6)
- Bravard, S. (2005). *Usages pédagogiques des QCM : un guide pour la mise en place d'un questionnaire à choix multiple* (Mémoire de master, Université de Poitiers, Poitiers). Repéré à l'URL http://fle.u-strasbg.fr/evaluation_fle/Bravard_qcm.pdf
- Brown, P. C., Roediger, H. L. et Mc Daniel, M. A. (2014). *Make it stick : the science of successful learning*, Harvard : Belknap Press.
- Butler, A. C., & Roediger, H. L. (2007). Testing improves long term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19(4-5), 514–527.
- Chan, C. (2009). Assessment: Short Answer Questions, Assessment Resources@HKU, University of Hong Kong. Repéré à l'URL <http://ar.cetl.hku.hk>
- Chua, Y. P., & Don, Z. M. (2013). Effects of computer-based educational achievement test on test performance and test takers' motivation. *Computers in Human Behavior*, 29(5), 1889–1895. <http://doi.org/10.1016/j.chb.2013.03.008>
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, 33(5), 593–602. <http://doi.org/10.1111/1467-8535.00294>
- Croisile, B. (2009). Approche neurocognitive de la mémoire. *Gérontologie et société*, (130), 11–29.
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The Testing Effect for Learning Principles and Procedures from Texts. *The Journal of Educational Research*, 107(5), 357–364. <https://doi.org/10.1080/00220671.2013.823370>
- Dolbec, A., & Clément, J. (2004). La recherche-action. Dans T. Karsenti, & L. Savoie-Zajc (Éds), *La recherche en éducation : étapes et approches* (2^e éd., pp. 181-208). Sherbrooke : CRP.
- Fernandez, J. (2016). Améliorer l'apprentissage à partir d'environnements informatiques: effets des quiz interactifs sur les processus cognitifs et métacognitifs. Repéré à l'URL https://www.researchgate.net/profile/Jonathan_Fernandez/publication/304014265_Ameliorer_l'Apprentissage_a_Partir_d'Environnements_Informatiques_Effets_des_Quiz_Interactifs_sur_les_Processus_Cognitifs_et_Metacognitifs/links/5763162b08aeab6e490aa2de.pdf
- Fournier, M. (1996, février-mars). La pédagogie différenciée. *Sciences humaines, Hors série* 12, 26.
- Glitzbach, M. (2017, 9 mars). Introducing the new Quizlet Learn [Billet de blogue]. Repéré à <https://quizlet.com/blog/introducing-the-new-quizlet-learn>
- Harmer, M. J. (2015). *The Practice of English Language Teaching 5th Edition Book with DVD Pack* (1st ed.). Pearson Longman.

Hattie, J. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (1st ed.). London ; New York: Routledge.

Hattie, J. (2011). *Visible Learning for Teachers: Maximizing Impact on Learning* (1st Edition). London ; New York: Routledge.

Leclercq, D. (1986). *La conception des Questions à Choix Multiple*. Bruxelles: Labor.

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., Mcdermott, K. B., & Roediger, H. L. (2011). Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement. *Journal of Educational Psychology*, 103(2), 399–414. <https://doi.org/10.1037/a0021782>

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, 19(4–5), 494–513. <https://doi.org/10.1080/09541440701326154>

McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in Middle-School Science: Successful Transfer Performance on Classroom Exams. *Applied Cognitive Psychology*, 27(3), 360–372. <https://doi.org/10.1002/acp.2914>

Metcalfe, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, 37(8), 1077–1087. <https://doi.org/10.3758/MC.37.8.1077>
Mühlemann, K., & Schöbi, N. (2016). Dépenses publiques d'éducation 2005-2014. Office fédéral de la statistique (OFS).

Nicolas, S. (1992). Hermann Ebbinghaus et l'étude expérimentale de la mémoire humaine. *L'année psychologique*, 92(4), 527–544. <https://doi.org/10.3406/psy.1992.29538>

Perrenoud, P. (1994). *Métier d'élève et sens du travail scolaire*. Paris: ESF éd.

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory : Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>

Roediger, H. L., & Karpicke, J. D. (2006b). Test-Enhanced Learning Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, 17(3), 249–255. <http://doi.org/10.1111/j.1467-9280.2006.01693.x>

Roediger, H. L., Putnam, A. L. et Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation*, 55.

Roediger III, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, 17(4), 382–395. <https://doi.org/10.1037/a0026252>

Seashore, R. H. (1948). The Importance of Vocabulary in Learning Language Skills. *Elementary English*, 25(3), 137–160.

Semb, G. B., & Ellis, J. A. (1994). Knowledge taught in school: What is remembered? *Review of Educational Research*, 64, 253–286.

Stefanou, C., & Révész, A. (2015). Direct Written Corrective Feedback, Learner Differences, and the Acquisition of Second Language Article Use for Generic and Specific Plural Reference. *The Modern Language Journal*, 99(2), 263–282. <https://doi.org/10.1111/modl.12212>

Taconnat, L. (2012). Fonctionnement et dysfonctionnement de la mémoire humaine. *Le Journal des psychologues*, (297), 18–23. <http://doi.org/10.3917/jdp.297.0018>

Thornbury, S. (2002). *How to Teach Vocabulary*. Longman.

Annexes

Annexe 1 : quiz 1 au format PDF	<i>4 pages</i>
Annexe 2 : quiz 2 au format PDF	<i>5 pages</i>
Annexe 3 : quiz 3 au format PDF	<i>4 pages</i>
Annexe 4 : quiz 4 au format PDF	<i>4 pages</i>
Annexe 5 : évaluation sommative	<i>6 pages</i>
Annexe 6 : résultats de la classe test sur toute la durée du dispositif	<i>1 page</i>
Annexe 7 : résultats des élèves pris en compte dans l'étude lors de l'évaluation sommative	<i>1 page</i>
Annexe 8 : calculs pour l'analyse de la première partie de l'étude	<i>3 pages</i>
Annexe 9 : scores en ligne des questions de l'évaluation finale	<i>2 pages</i>

Annexe 1

Quiz 1 au format PDF

Review Quizz 1



répondre en faisant de ton mieux.

Ce quizz contient les éléments vus durant la semaine écoulée. Essaie d'y

Q.1) Modalverben

In einem Zug [Blank] man nicht rauchen!



Explanation : Le verbe de modalité correct ici est: dürfen = avoir la permission. nicht dürfen = être interdit. On le conjugue avec la 3e personne du présent.

Q.2) In der Schule

Vor dem Lehrerzimmer gehst du [Blank].



Explanation : gauche = links (pense à l'anglais: left)

Q.3) In der Schule

Was ist das?



- A. eine Turnhalle
- B. eine Sporthalle
- C. eine Sportsaal
- D. ein Sportzimmer

Explanation: Das ist eine Sporthalle. Hier kann man fast alle Sportarten machen.

Q.4) Modalverben
Danke für die Einladung! Was _____ ich mitbringen?

- A. darf
- B. will
- C. muss
- D. kann

Explanation: Le verbe de modalité qui convient ici est: können = pouvoir. Que puis-je apporter?

Q.5) In der Schule
Hier gibt es Frau Casciotta, Frau Maire, Frau Stauffer, ein Telefon, ein Fax, 4 Computer und viele Informationen.

- A. die Cafeteria
- B. der Computerraum
- C. das Sekretariat
- D. die Direktion

Explanation: Frau Casciotta, Frau Maire und Frau Stauffer sind 3 Sekretärinnen in der Schule. Sie arbeiten im Sekretariat.

Q.6) Modalverben
Andrea und Marta _____ sehr gut Deutsch sprechen.

- A. kann
- B. darf
- C. können
- D. dürfen

Explanation: Le verbe de modalité qui convient ici est können = savoir. Il est conjugué à la 1ère personne du pluriel (wir).

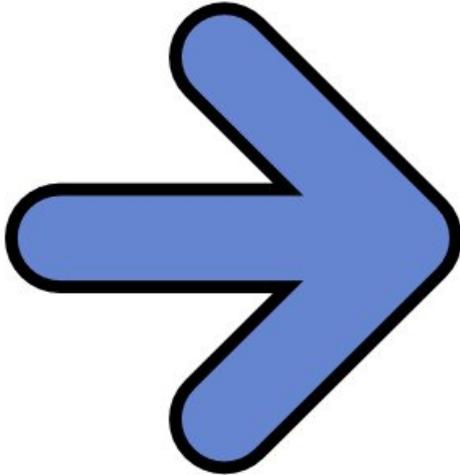
Q.7) In der Schule
Hier gibt es oft Experimente: [Blank]

Explanation : Im Chemieraum macht man Experimente.

Q.8) Modalverben
Heute Abend [Blank] ich bis 23 Uhr mit meinen Freundinnen ausgehen. Meine Eltern sind ok.

Explanation : Le verbe de modalité approprié ici est: dürfen = avoir la permission. On le conjugue à la 3e personne du présent.

Q.9) In der Schule
Dann gehst du [Blank] und du findest die Toilette!



Explanation : à droite = rechts. Pense à l'anglais: right

Q.10) Modalverben
Herr Matthey, [Blank] ich trinken gehen, bitte?

Explanation : Le verbe de modalité à utiliser ici est: dürfen = avoir la permission. On le conjugue à la 1ère personne du singulier.

Q.11) In der Schule

Wie heisst dieses Zimmer?



- A. der Computerraum
- B. der Computeraum
- C. das Computerzimmer
- D. das Lehrerzimmer

Explanation: Das ist der Computerraum, mit 2 'r'! Computer + Raum = Computerraum!

Q.12) Modalverben

Herr Matthey, _____ Sie bitte langsamer sprechen? Ich verstehe nicht!

- A. können
- B. wollen
- C. müssen
- D. dürfen

Explanation: Le verbe de modalité approprié ici est: können = pouvoir. On le conjugue avec le Sie de politesse.

Q.13) In der Schule

In einer Schule, wie heisst dieses Zimmer?



- A. der Hausmeister
- B. der Erdgeschoss
- C. die Mensa
- D. die Bibliothek

Explanation: In einer Bibliothek findet man viele Bücher!

Q.14) In der Schule

Hier nach dem Sekretariat gehst du [Blank].



Explanation : Tout droit = geradeaus. / die Gerade = la ligne droite

Q.15) In der Schule

Hier sitzen die Experten für Bio, Mathe, Englisch, Französisch, usw.

- A. der Chemieraum
- B. der Computerraum
- C. der Schulhof
- D. das Lehrerzimmer

Explanation: Im Lehrerzimmer sitzen alle Lehrer von der Schule.

Q.16) In der Schule

Hier sitzt der Chef (Herr Jubin): [Blank]

Explanation : In Le Locle, der Direktor ist Herr Jubin. Er arbeitet in der Direktion.

Annexe 2

Quiz 2 au format PDF

Review quizz 2



Bienvenue sur ce 2e quizz!

Tu trouveras ci-après des questions en lien avec la matière vue en classe la semaine écoulée. Réponds aux questions du mieux possible et seul.

Bon quizz!

Max Time Allowed: 10 min

Q.1) Wortschatz

Ecris ce que tu vois sur cette image dans la lacune. La description ci-dessous peut t'aider:
Hier kommen jeden Tag 300 Schüler und Schülerinnen: [Blank]



Q.2) das Zimmer

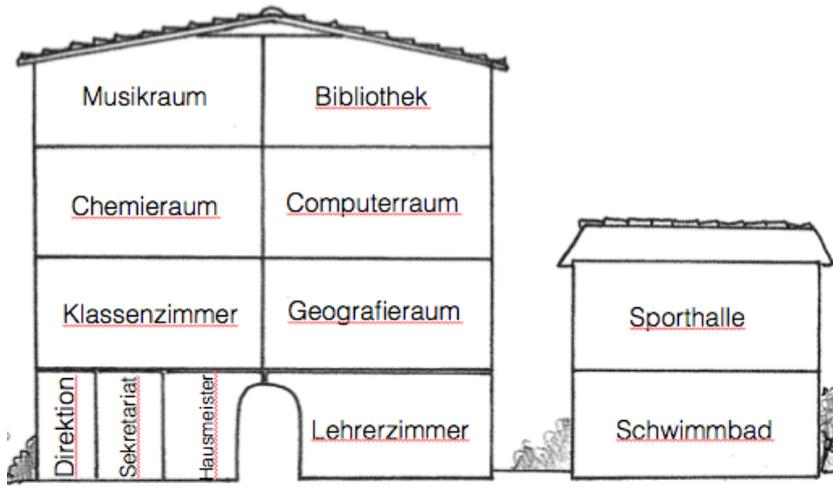
Quelle est la phrase incorrecte? Choisis parmi les 4 possibilités:



- A. Auf dem Sofa sitzt ein Teddybär.
- B. Vor dem Computer steht ein Stuhl.
- C. Auf dem Boden liegt ein Teppich.
- D. Das Regal steht über dem Fenster.

Q.6) In der Schule

Regarde l'image et choisis la description qui n'est pas correcte.



- A. das Lehrerzimmer ist gleich am Eingang rechts.
- B. der Musikraum ist links neben der Bibliothek.
- C. der Musikraum ist im dritten Stock links.
- D. das Schwimmbad ist im ersten Stock.

Q.7) Verben

Choisis parmi la liste la phrase qui décrit le mieux l'image. Attention à la conjugaison!



- A. die Gitarre liegt an der Wand.
- B. die Gitarre hängt an der Wand.
- C. die Gitarre steht an der Wand.
- D. die Gitarre liegen an der Wand.
- E. die Gitarre stehen an der Wand.
- F. die Gitarre hängen an der Wand.

Q.8) Wortschatz

Qu'est-ce qui est représenté sur cette image? Choisis parmi les propositions:



- A. die Treppe
- B. der Schulhof
- C. die Schultasche
- D. der Stock

Q.9) Wortschatz

Qui est représentée sur cette image? Choisis parmi les propositions:



- A. die Lehrerin
- B. die Hausmeisterin
- C. die Mehrerin
- D. die Lekehrin

Q.10) Präpositionen

Wo ist der Mann? Remplis la lacune avec la préposition la plus indiquée: der Mann ist [Blank] dem Auto.



Q.11) Präpositionen

Wo ist der Mann? Choisis la préposition qui correspond le plus à la situation.



- A. zwischen
- B. an
- C. auf
- D. in

Q.12) Präpositionen

Wo ist der Mann? Choisis la préposition qui correspond le plus à la situation.

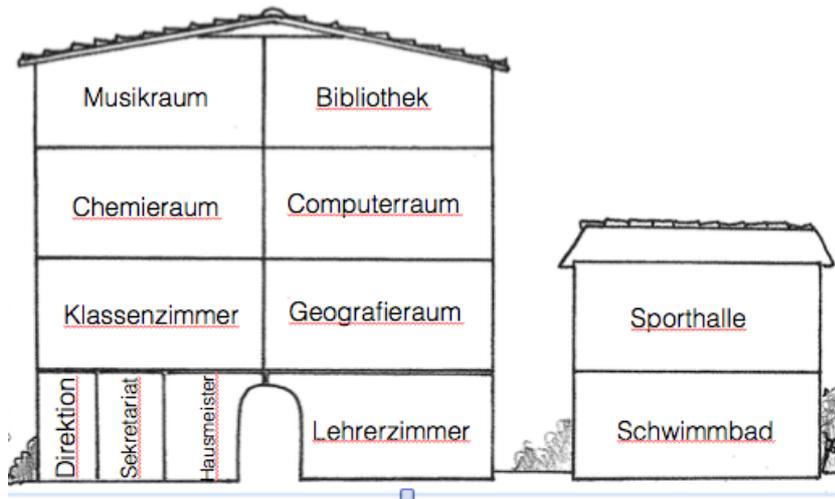


- A. in
- B. hinter
- C. über
- D. unter

Q.13) In der Schule

Complète les lacunes en fonction de l'image. La 2e lacune est la fin du déterminant.

Der Computerraum ist im zweiten [Blank] rechts, neben d[Blank] Chemieraum.



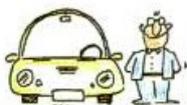
Q.14) Präpositionen

Wo ist der Mann? Remplis la lacune avec la préposition la plus indiquée: Der Mann steht [Blank] dem Auto.



Q.15) Präpositionen

Wo ist der Mann? Remplis la lacune avec la préposition la plus indiquée. Attention, l'homme ne touche pas la voiture! Der Mann steht [Blank] dem Auto.



Annexe 3

Quiz 3 au format PDF

Review quizz 3



Bienvenue sur ce 3e quizz!

Tu trouveras ci-après des questions en lien avec la matière vue en classe la semaine écoulée. Réponds aux questions du mieux possible et seul.

Bon quizz!

Q.1) Correction:

Choisis parmi les propositions celle qui corrige toutes les erreurs de la phrase ci-dessous:

Das Bild steht auf dem Boden.



- A. Das Bild hängt auf dem Boden.
- B. Das Bild steht an der Wand.
- C. Das Bild hängt an dem Boden
- D. Das Bild hängt an der Wand.

Explanation: La bonne réponse est: "das Bild hängt an der Wand". hängen = être pendu / die Wand = le mur, la paroi

Q.2) Descriptions

Voici le plan d'une école. Indique quelle pièce correspond à la description ci-dessous (déterminant + nom: ex: der Chemieraum):

Sie ist rechts neben dem Sekretariat.

L I F T	Kunstraum	Klasse 308	T R E P P E	Musikraum
	Sekretariat	Direktion		Cafeteria
	Chemieraum	Hausmeister		Bibliothek
Pausenhof				

Explanation : Il s'agit de la pièce qui se trouve au 1er étage au centre: die Direktion = la direction.

Q.3) Descriptions

Voici le plan d'une école. Indique quelle pièce correspond à la description ci-dessous (déterminant + nom: ex: der Chemieraum):

Er ist im zweiten Stock ganz rechts.

L I F T	Kunstraum	Klasse 308	T R E P P E	Musikraum
	Sekretariat	Rektorat		Cafeteria
	Sporthalle	Hausmeister		Bibliothek
Pausenhof				

Explanation : Il s'agit de la pièce qui se trouve au 2e étage tout à droite: der Musikraum = la salle de musique.

Q.4) Correction:

Regarde l'image. Réécris ensuite la phrase ci-dessous en apportant la/les correction(s). Ne mets pas de point à la fin:

Turbo sitzt neben dem Computer:

[Blank]



Explanation : Turbo sitzt nicht, sie steht! Stehen = être debout

Q.5) Directions

Complète le texte ci-dessous avec les bonnes prépositions ou indications de direction. Attention, tu dois parfois mettre la préposition et le déterminant (ex: neben dem):

Die Direktion ist [Blank] Cafeteria. Die Toiletten? Gehen sie immer [Blank].

Der Computerraum ist [Blank] Klasse 7a.

Cafeteria	Toilette	Computerraum	Schwimmhalle
Direktion		7a	
Sekretariat		Lehrzimmer	
Hausmeister			

Explanation : La direction est devant la cafétéria = vor der. Les toilettes sont tout droit = geradeaus. La salle d'informatique est derrière la classe 7a = hinter der. Vor et hinter demandent le Dativ, c'est pour cela que 'die Cafeteria' devient 'vor der Cafeteria' et 'die Klasse 7a' devient 'hinter der Klasse 7a'.

Q.6) Directions

Choisis parmi les options proposées celles qui permettent de compléter le ci-dessous. Chaque terme séparé par des '/' correspond à une lacune:

Die Cafeteria finden sie im Erdgeschoss _____. Das Lehrerzimmer ist am Eingang _____. Der Hausmeister, Die Direktion und das Sekretariat sind _____. Die Schwimmhalle ist _____ der Schule.



- A. hinten links / geradeaus / links / zwischen
- B. hinten links / rechts / links / neben
- C. vorne links / links / rechts / neben
- D. hinten links / links / rechts / neben

Explanation: Hinten = derrière / vorne = devant / links = à gauche / rechts = à droite / neben = à côté / zwischen = entre

Q.7) Wortschatz:

Fais correspondre chaque mot avec sa définition:

	Choice		Match
A	Sie arbeitet in einer Schule mit Schülern:	1	der Hausmeister
B	Dort kann man laufen und Sport treiben.	2	die Sporthalle
C	Er putzt die Schule und arbeitet auch am Abend.	3	die Treppe
D	Dort kann man etwas zum Trinken kaufen.	4	die Lehrerin
E	Man kann sie hoch oder runter gehen	5	die Cafeteria

Q.8) Traduction:

Choisis la traduction qui convient pour la phrase ci-dessous:

Au rez-de-chaussée.

- A. der Erdgeschoss
- B. in Erdgeschoss
- C. Im Geschoss
- D. Im Erdgeschoss

Explanation: La réponse correcte est: im Erdgeschoss. im = in dem => in dem Erdgeschoss

Q.9) Traduction:

Choisis la traduction qui convient pour la phrase ci-dessous:

Du gehst die Treppe hoch.

- A. Tu passes par la trappe du haut.
- B. Tu descends les escaliers.
- C. Tu prends l'ascenseur.
- D. Tu montes les escaliers.

Explanation: La réponse correcte est: du gehst die Treppe hoch. Hoch = en-haut => monter

Q.10) Präpositionen:

Regarde l'image puis complète la lacune avec la préposition qui convient, pour décrire le numéro 1 dans l'image:

Die Blumen (les fleurs) stehen [Blank] dem Tisch.



Explanation : La préposition qui convient ici est: auf = sur

Q.11) Präpositionen:

Regarde l'image puis complète la lacune avec la préposition qui convient, pour décrire le numéro 7 dans l'image:

Das Bett steht [Blank] dem Schreibtisch.



Explanation : La préposition qui convient ici est: neben = à côté. On peut y ajouter la précision 'links' ou 'rechts' pour savoir de quel côté.

Q.12) Traduction:

Traduis la phrase ci-dessous, en faisant attention aux majuscules. Ne mets pas de point à la fin:

Tu descends les escaliers.
[Blank]

Explanation : La traduction correcte est: du gehst die Treppe runter. runter = en bas / die Treppe = les escaliers

Q.13) Traduction:

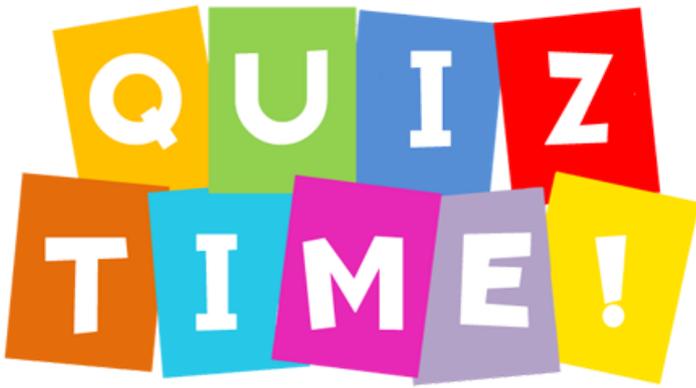
Traduis la phrase ci-dessous, en faisant attention aux majuscules. Ne mets pas de point à la fin:

Au premier étage.
[Blank]

Annexe 4

Quiz 4 au format PDF

Review quizz 4



Bienvenue sur ce 4e quizz!

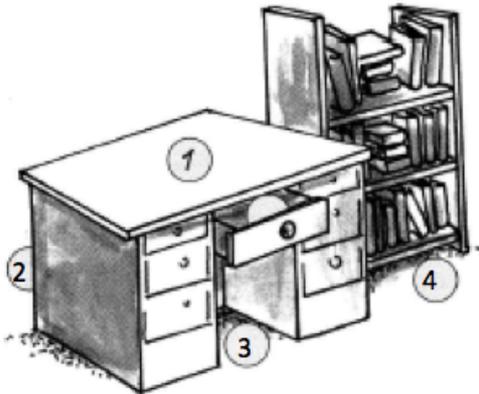
Tu trouveras ci-après des questions en lien avec la matière vue en classe la semaine écoulée. Réponds aux questions du mieux possible et seul.

Bon quizz!

Q.1) Die Klassenarbeit

Décris le numéro 4 de l'image. Utilise une préposition pour la 1ère lacune et un déterminant pour la 2e lacune.

Wo sucht Herr Schmidt die Klassenarbeit? [Blank] [Blank] Bücherregal.



Explanation : La préposition est 'vor' (devant) ou 'unter' (sous, en-dessous- de). Le déterminant est 'dem', car il s'agit du Dativ de 'das' (das Bücherregal).

Q.2) In der Schule

Complète le dialogue avec 2 mots (1 par lacune) parmi la liste ci-dessous:

Entschuldigung - finde - im - neben - runter - Stock - suche - wo

- Wo [Blank] ich die Toilette, bitte?

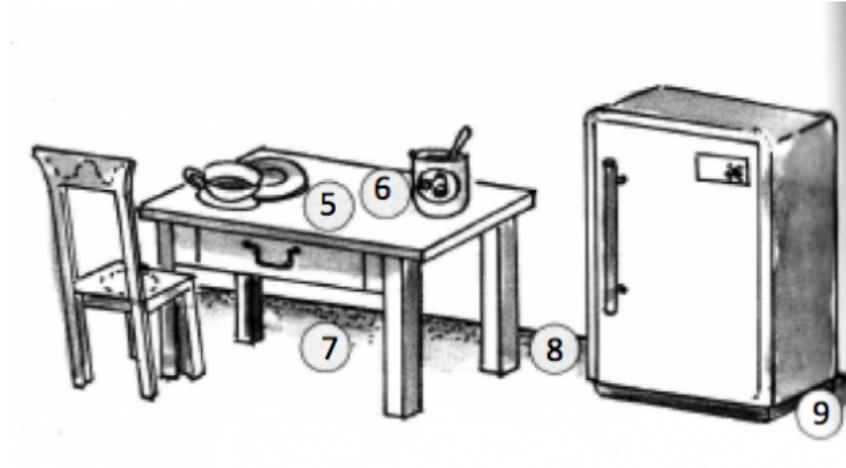
- Hier die Treppe [Blank], im Erdgeschoss.

Explanation : Wo finde ich die Toilette? = où puis-je trouver les toilettes? finden = trouver. die Treppe runter = en bas les escaliers

Q.3) Die Klassenarbeit

Décris le numéro 9 de l'image. Utilise une préposition pour la 1ère lacune et un déterminant pour la 2e lacune.

Wo sucht Herr Schmidt die Klassenarbeit? [Blank] [Blank] Kühlschrank.

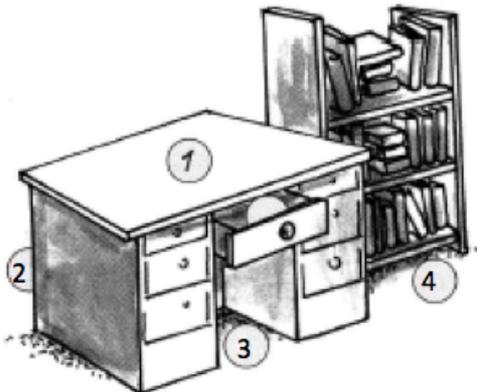


Explanation : La préposition est 'neben' (à côté). Le déterminant est 'dem', car il s'agit du Dativ de 'der' (der Kühlschrank).

Q.4) Die Klassenarbeit

Décris le numéro 2 de l'image à l'aide d'une des propositions.

Wo sucht Herr Schmidt?



- A. Auf die Wand.
- B. An die Wand.
- C. An der Wand.
- D. Auf der Wand.

Explanation: La préposition correct est 'an' (contre) et le déterminant 'der', car c'est le Dativ de 'die' (die Wand).

Q.5) In der Schule

Choisis parmi les propositions celles qui permettent de compléter le dialogue correctement.

- _____, wo ist das Lehrerzimmer, bitte?
- Die Treppe _____, im ersten _____ links.

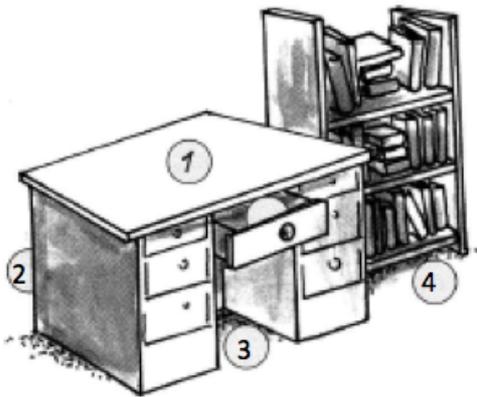
- A. Entschuldigung / hoch / Stock
- B. Stock / hoch / Klasse
- C. Stock / hoch / Lift
- D. Entschuldigung / Stock / geradeaus

Explanation: La réponse correcte est: Entschuldigung, wo ist das Lehrerzimmer, bitte? (Excusez-moi, où se trouve...). im ersten Stock = au premier étage.

Q.6) Die Klassenarbeit

Décris le numéro 3 de l'image à l'aide d'une des propositions.

Wo sucht Herr Schmidt?



- A. Neben dem Schreibtisch.
- B. Unter dem Regal.
- C. Über dem Schreibtisch.
- D. Unter dem Schreibtisch.

Explanation: La préposition correct est 'unter' (sous) et le déterminant 'dem', car c'est le Dativ de 'der' (der Schreibtisch).

Q.7) Beschreibungen

Parmi les propositions ci-dessous, choisis celle qui décrit le mieux l'image ci-contre:



- A. das Bild sitzt an der Wand.
- B. das Bild liegt an der Wand.
- C. das Bild steht an der Wand.
- D. das Bild hängt an der Wand.

Explanation: Le verbe correct est 'hängen' = être suspendu

Q.8) Beschreibungen

Complète les lacunes de la phrase ci-dessous, pour décrire l'image. La première lacune est un verbe de position, la 2e une préposition.

Die Schüler [Blank] [Blank] der Schule.

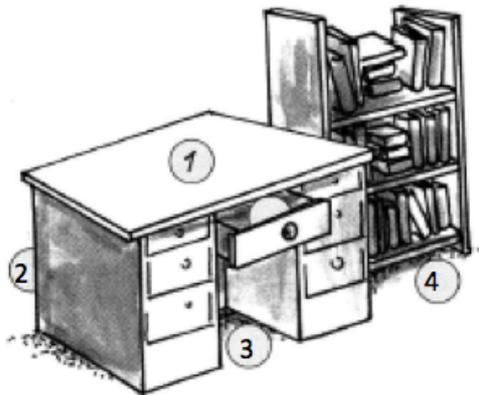


Explanation : stehen = être debout. Le verbe est conjugué au pluriel car il y a plusieurs élèves (die Schüler). vor = devant. Les élèves se trouvent devant l'école. Ev. neben aussi correct.

Q.9) Die Klassenarbeit

Décris le numéro 1 de l'image. Utilise une préposition pour la 1ère lacune et un déterminant pour la 2e lacune.

Wo sucht Herr Schmidt die Klassenarbeit? [Blank] [Blank] Schreibtisch.



Explanation : La préposition est 'auf' (sur). Le déterminant est 'dem', car il s'agit du Dativ de 'der' (der Schreibtisch).

Annexe 5

Evaluation sommative

Test Geni@l A1 – Einheit 12

Note :

*Temps à disposition: 45'***Directions**

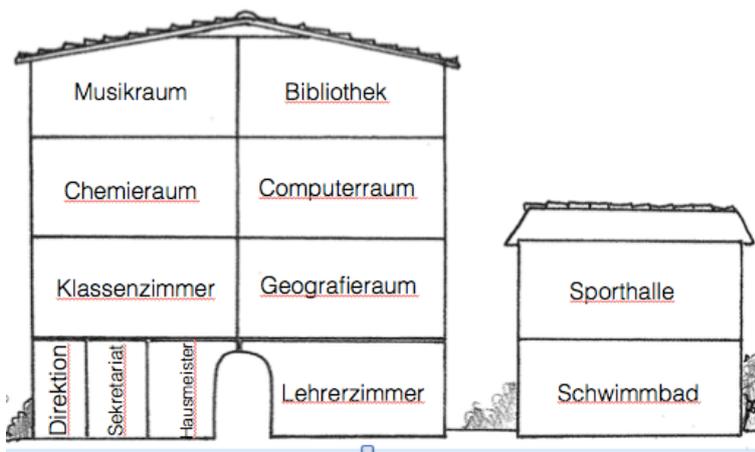
- Voici le plan d'une école. Indique quelle pièce correspond à chaque description.

- N'oublie pas le déterminant !

L I F T	Kunstraum	Klasse 308	T R E P P E	Musikraum
	Sekretariat	Direktion		Cafeteria
	Chemieraum	Hausmeister		Bibliothek
Pausenhof				

- 1) Sie ist rechts neben dem Sekretariat : _____
- 2) _____ findest du im zweiten Stock ganz links.

- Complète les lacunes en fonction de l'image. La 2e lacune est la fin du déterminant.



- 3) Der Computerraum ist im zweiten _____ rechts, neben d_____ Chemieraum.

- Toujours selon l'image, complète la lacune avec la bonne préposition :

- 4) Der Musikraum ist im dritten Stock, _____ dem Chemieraum.

Traductions

- Traduis la phrase suivante en français:

5) Du gehst die Treppe hoch. _____

- Traduis les phrases ci-dessous en allemand :

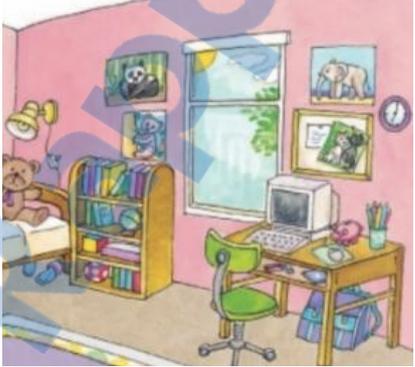
6) Au premier étage. _____

7) Au rez-de-chaussée _____

8) Les élèves demandent. _____

Corrections

- Regarde les images et réécris les phrases de droite en apportant la correction nécessaire (1 erreur par phrase).

	<p>9) Turbo liegt neben dem Computer.</p> <p>_____</p>
	<p>10) Turbo hängt auf dem Stuhl.</p> <p>_____</p>
	<p>11) Das Regal steht über dem Fenster.</p> <p>_____</p> <p>12) Die Poster hängen an der Tür.</p> <p>_____</p>
	<p>13) Der Mann steht unter dem Auto.</p> <p>_____</p>

	<p>14) das Schwimmbad ist im ersten Stock.</p> <hr/> <p>15) Das Sekretariat ist links von dem Direktion.</p> <hr/>
--	--

Vocabulaire

- *Quels mots du vocabulaire sont représentés par ces images.*
- *Indique ta réponse dans la colonne de droite, en précisant le déterminant.*



16)



17)



18)



19)



20)

- Complète la lacune avec la direction indiquée par la flèche :



21) Hier nach dem Sekretariat gehst du _____.

- De quelle pièce de l'école s'agit-il ? Répond en indiquant le déterminant :

22) Hier gibt es oft Experimente: _____

- De qui s'agit-il ? Indique le déterminant.

23) Er arbeitet in einer Schule mit Schülern. Er ist ein Spezialist für Mathe, Deutsch, Geschichte, usw. _____

- De quoi s'agit-il ? N'oublie pas le déterminant.

24) Man kann sie hoch oder runter gehen : _____

Verbes de modalité

- Complète la lacune avec le verbe de modalité correct, conjugué à la bonne personne :



25) In einem Zug _____ man nicht rauchen!

26) Herr Matthey, _____ Sie bitte langsamer sprechen? Ich verstehe nicht!

27) Andrea und Marta _____ sehr gut Deutsch sprechen.

28) Die Schüler _____ ihre Hausaufgaben machen.

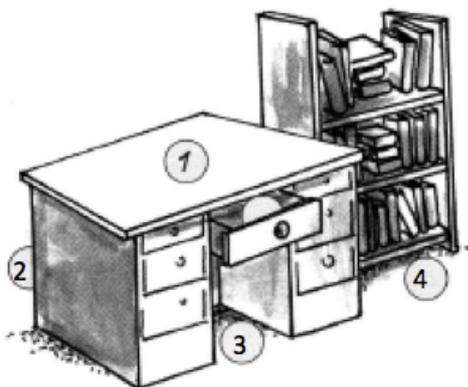


29) _____ du noch ein Glas Wein trinken ? – Ja, gern.

Prépositions

- Wo ist der Mann? Remplis la lacune avec la préposition la plus indiquée:

	<p>30) der Mann ist _____ dem Auto.</p>
	<p>31) Der Mann sitzt _____ dem Auto.</p>
	<p>32) Der Mann steht _____ dem Auto.</p>
	<p>33) Der Mann steht _____ den Autos.</p>
	<p>34) Der Mann steht _____ dem Auto.</p>



- Décris le numéro 3 de l'image à l'aide d'une préposition.

35) Wo sucht Herr Schmidt? _____ dem Schreibtisch.

- Décris le numéro 2 de l'image à l'aide d'une des préposition :

36) Wo sucht Herr Schmidt ? _____ der Wand.

Verbes de position

- Remplis les lacune avec le verbe qui décrit le mieux l'image. Attention à la conjugaison!

37) Wie sind die Bücher? Die Bücher _____ auf dem Regal.



38) Das Bild _____ an der Wand.



- Complète les lacunes de la phrase ci-dessous, pour décrire l'image.
 - La première lacune est un verbe de position, la 2e une préposition.

39) Die Schüler _____ _____ der Schule.



Annexe 6

Résultats de la classe test sur toute la durée du dispositif

Résultats de la classe test sur toute la durée du dispositif

Elève	Moyenne au 10.02	Score Quiz 1	Score Quiz 2	Score Quiz 3	Score Quiz 4	Score review Quiz	Evaluation finale	Note
E1	5.95	37.5	33	38	85.71	Dû recommencer	0.897435897	6
E2	4.23	19.05	-	22.86	14.29	25.95	0.397435897	3
E3	4.77	50	27.78	48.57	28.57	49.62	0.730769231	5
E4	5	19.05	16.67	20	78.57	56.49	0.666666667	4.5
E5	4.59	37.5	22.22	37.14	35.71	51.15	0.525641026	3.5
E6	5.27	42.86	61.11	54.29	57.14	60.31	0.858974359	5.5
E7	2.73	-	33.33	37.14	42.86	Dû recommencer et pas pu finir	0.730769231	5
E8	3.73	48.86	55.56	-	57.14	-	0.455128205	3.5
E9	3.36	14.29	13	0	7.14	18.32	0.262820513	2
E10	3.86	23.81	0	22.86	0	Pas pu finir	0.384615385	3
E11	4.36	23.81	33.33	28.57	28.57	36.64	0.641025641	4.5
E12	5.32	31.25	66.67	54.29	85.71	70.99	0.833333333	5.5
E13	2.42	6.25	-	11.43	14.29	15.27	0.185897436	2
E14	3.91	-	27	0	-	-	0.211538462	2
E15	5.36	37.5	50	71.43	-	60.31	0.903846154	6
E16	4.45	50	27	46	50	58.78	0.647435897	4.5
E17	3.36	31.25	22.22	34.29	42.86	Dû recommencer	0.461538462	3.5
E18	4	38.1	16.67	48.57	14.29	43.51	0.576923077	4
E19	4.83	56.25	44.44	-	71.43	55.73	0.698717949	4.5
E20	2.6	28.57	22.22	-	-	-	-	
Effectif	20	18	18	17	17	17	19	
Moyenne	4.205	33.1055556	31.79	33.8494118	42.0164706	46.39	0.58265857	
Ecart-type	0.953176269	13.3403624	17.1541113	18.9790127	26.8235822	16.74714855	0.219883989	

Annexe 7

Résultats des élèves pris en compte dans l'étude lors de l'évaluation sommative

Résultats des élèves pris en compte dans l'étude à l'évaluation sommative

Question	Mode de test initial	Ordre de présentation	E18	E4	E12	E9	E3	E6	E16	E5	E11	E17	E1	Score ligne
1	RC	1	1	0.5	0.5	1	0.5	1	1	0.5	1	0.5	0.5	0.73
33	Non testé	2	0	0	0	0	0.5	1	1	0	0	0	1	0.32
12	RC	3	0.5	1	0.75	1	1	1	0.5	1	0.75	1	1	0.86
29	Non testé	4	1	0	1	0	1	0	1	1	1	0	1	0.64
2	RC	5	0.5	1	1	1	1	0.75	1	1	1	1	1	0.93
3	RC	6	0.75	1	0.75	0.75	1	1	1	1	1	0.5	1	0.89
23	QCM	7	0.25	0.75	0.75	0.25	1	1	1	0.5	0.75	0.75	0.75	0.70
36	Non testé	8	0.5	1	1	0.5	0.5	1	0.75	0.75	0.5	0.75	1	0.75
4	RC	9	1	1	1	0.5	1	1	1	1	1	1	1	0.95
35	Non testé	10	1	0	1	0	0.75	1	0.75	1	1	1	0.75	0.75
25	QCM	11	0	1	1	0	1	1	1	1	1	0	1	0.73
37	Non testé	12	0	0.75	0	0	1	1	1	0	1	0	1	0.52
30	Non testé	13	1	1	1	0.5	1	1	1	1	1	1	1	0.95
16	QCM	14	0	1	1	0	1	1	0.75	1	0.75	0.5	1	0.73
38	Non testé	15	1	1	1	0.5	1	0.5	1	0.75	0	0	1	0.70
8	RC	16	1	1	1	0.25	1	1	1	1	0.75	0.25	1	0.84
9	RC	17	0.5	0.5	1	0.25	1	1	1	1	1	0.75	1	0.82
19	QCM	18	0.25	0.25	0.5	0	0.75	0.75	1	0	0.5	1	1	0.55
22	QCM	19	1	0.5	1	0.25	1	1	0.75	1	1	0.5	1	0.82
39	Non testé	20	0.25	1	1	0.5	1	1	0.5	0	0.75	1	1	0.73
5	RC	21	0	0.75	1	0	0	1	0	0	0.75	1	1	0.50
6	RC	22	0.25	0.25	0.75	0	0.75	0.5	0.25	0	0.25	1	1	0.45
34	Non testé	23	0.5	1	0.5	0	0	1	0.5	1	0	0	0.5	0.45
24	QCM	24	1	1	1	0	1	1	1	0	0	0	1	0.64
7	RC	25	0	0	1	0	0	0	0	0	0	0	1	0.18
20	QCM	26	0	0.5	1	0	0	0.5	0	0	0.5	0	1	0.32
21	QCM	27	1	0	1	0.5	0.5	1	0.5	0	0	0	1	0.50
31	Non testé	28	0	0	1	0	0	1	0	0	0.5	0	0	0.23
32	Non testé	29	0	0	0.5	0	0	0	0	0	0	0	1	0.14
10	RC	30	1	1	1	0	1	1	1	1	1	0	1	0.82
17	QCM	31	1	1	1	0	1	1	0.75	1	1	0	1	0.80
18	QCM	32	1	1	1	0	1	1	0	0	0	0	1	0.55
27	Non testé	33	0	1	0.75	0	1	1	0	0	1	0.75	0	0.50
28	Non testé	34	1	1	1	0	1	1	0	0	0	1	1	0.64
15	QCM	35	1	1	1	0.75	1	1	1	1	1	0.75	1	0.95
26	QCM	36	1	0	0	0	0	1	0	0	1	0	1	0.36
11	RC	37	1	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.59
14	QCM	38	0.75	1	0.75	0.75	1	1	1	1	1	1	1	0.93
13	RC	39	0.5	0.75	1	0.5	0.75	1	0.75	0.5	0.75	0.5	1	0.73
			22.5	26	32.5	10.25	28.5	33.5	25.25	20.5	25	18	35	
			57.69%	66.67%	83.33%	26.28%	73.08%	85.90%	64.74%	52.56%	64.10%	46.15%	89.74%	

Annexe 8

Calculs pour l'analyse de la première partie de l'étude

Calculs pour l'analyse de la première partie de l'étude

	E18	E4	E12	E9	E3	E6	E16	E5	E11	E17	E1
Résultat aux questions préalablement testées au format QCM	63.46%	69.23%	84.62%	19.23%	78.85%	94.23%	67.31%	50.00%	65.38%	34.62%	98.08%
Résultat aux questions préalablement testées au format réponses courtes	61.54%	71.15%	90.38%	44.23%	73.08%	82.69%	69.23%	65.38%	75.00%	61.54%	92.31%
Résultat aux questions non testées	48.08%	59.62%	75.00%	15.38%	67.31%	80.77%	57.69%	42.31%	51.92%	42.31%	78.85%

Moyenne QCM	65.91%
Moyenne RC	71.50%
Moyenne non testé	56.29%

Analyse du testing effect

Résultats QCM	Résultats RC	Résultats non testés
0.634615385	0.615384615	0.480769231
0.692307692	0.711538462	0.596153846
0.846153846	0.903846154	0.75
0.192307692	0.442307692	0.153846154
0.788461538	0.730769231	0.673076923
0.942307692	0.826923077	0.807692308
0.673076923	0.692307692	0.576923077
0.5	0.653846154	0.423076923
0.653846154	0.75	0.519230769
0.346153846	0.615384615	0.423076923
0.980769231	0.923076923	0.788461538
0.307692308	0.538461538	0.307692308

Analyse de validité statistique

Analyse avec Friedman rank sum test

Statistique observée

Qobs: 15.36170213

p-value : 0.000461582

Paramètre du test: 2

Analyse de ceux qui ont progressé, stagné, ou régressé

variable 1 (moyenne), variable 2 (note au TE)

Moyenne au 10.02	Note évaluation sommative	Différence
5.95	6	0.84%
4.77	5	4.82%
5.00	4.5	-10.00%
4.59	3.5	-23.75%
5.27	5.5	4.36%
3.36	2	-40.48%
4.36	4.5	3.21%
5.32	5.5	3.38%
4.45	4.5	1.12%
3.36	3.5	4.17%
4.00	4	0.00%

Nb Progressé	7	3.13%
Nb Stagné	1	
Nb Régressé	4	-24.74%

Analyse avec Student test

Méthode : Paired t-test; Alternative :two.sided

Statistique observée Qobs : 1.039888198696

p-value : 0.32288122576875

Moyenne : 0.1755 Intervalle de confiance à 95%[-0.2005 ; 0.5514]

Degrés de liberté : 10

Moyenne des differences : 0.17545454545455

Voir quel type d'élève est favorisé par le dispositif

<i>Médiane des moyennes = 4.59</i>					
Groupe 1 (inférieur à la médiane)			Groupe 2 (supérieur à la médiane)		
<i>Moyenne</i>	<i>Note</i>	<i>Progression</i>	<i>Moyenne</i>	<i>Note</i>	<i>Progression</i>
3.36	2	<	5.95	6	>
3.36	3.5	>	5.32	5.5	>
4.00	4	=	5.27	5.5	>
4.36	4.5	>	5.00	4.5	<
4.45	4.5	>	4.77	5	>

3 sur 5 ont amélioré

4 sur 5 ont amélioré

4.59	3.5	<
------	-----	---

p =

0.31731050786291

donc pas statistiquement significatif

p = 0.179712494879

donc pas statistiquement significatif

Moyenne groupe 1 ayant progressé

2.83%

Moyenne groupe 2 ayant progressé

3.35%

Annexe 9

Scores en ligne des questions de l'évaluation finale

Scores des questions de l'évaluation sommative, par ligne

Question	Mode de test initial	Ordre de présentation	E18	E4	E12	E9	E3	E6	E16	E5	E11	E17	E1	Score ligne
33	Non testé	2	0	0	0	0	0.5	1	1	0	0	0	1	0.33
29	Non testé	4	1	0	1	0	1	0	1	1	1	0	1	0.64
36	Non testé	8	0.5	1	1	0.5	0.5	1	0.75	0.75	0.5	0.75	1	0.75
35	Non testé	10	1	0	1	0	0.75	1	0.75	1	1	1	0.75	0.75
37	Non testé	12	0	0.75	0	0	1	1	1	0	1	0	1	0.52
30	Non testé	13	1	1	1	0.5	1	1	1	1	1	1	1	0.95
38	Non testé	15	1	1	1	0.5	1	0.5	1	0.75	0	0	1	0.70
39	Non testé	20	0.25	1	1	0.5	1	1	0.5	0	0.75	1	1	0.73
34	Non testé	23	0.5	1	0.5	0	0	1	0.5	1	0	0	0.5	0.45
31	Non testé	28	0	0	1	0	0	1	0	0	0.5	0	0	0.23
32	Non testé	29	0	0	0.5	0	0	0	0	0	0	0	1	0.14
27	Non testé	33	0	1	0.75	0	1	1	0	0	1	0.75	0	0.50
28	Non testé	34	1	1	1	0	1	1	0	0	0	1	1	0.64
23	QCM	7	0.25	0.75	0.75	0.25	1	1	1	0.5	0.75	0.75	0.75	0.70
25	QCM	11	0	1	1	0	1	1	1	1	1	0	1	0.73
16	QCM	14	0	1	1	0	1	1	0.75	1	0.75	0.5	1	0.73
19	QCM	18	0.25	0.25	0.5	0	0.75	0.75	1	0	0.5	1	1	0.55
22	QCM	19	1	0.5	1	0.25	1	1	0.75	1	1	0.5	1	0.82
24	QCM	24	1	1	1	0	1	1	1	0	0	0	1	0.64
20	QCM	26	0	0.5	1	0	0	0.5	0	0	0.5	0	1	0.32
21	QCM	27	1	0	1	0.5	0.5	1	0.5	0	0	0	1	0.50
17	QCM	31	1	1	1	0	1	1	0.75	1	1	0	1	0.80
18	QCM	32	1	1	1	0	1	1	0	0	0	0	1	0.55
15	QCM	35	1	1	1	0.75	1	1	1	1	1	0.75	1	0.95
26	QCM	36	1	0	0	0	0	1	0	0	1	0	1	0.36
14	QCM	38	0.75	1	0.75	0.75	1	1	1	1	1	1	1	0.93
1	RC	1	1	0.5	0.5	1	0.5	1	1	0.5	1	0.5	0.5	0.73
12	RC	3	0.5	1	0.75	1	1	1	0.5	1	0.75	1	1	0.86
2	RC	5	0.5	1	1	1	1	0.75	1	1	1	1	1	0.93
3	RC	6	0.75	1	0.75	0.75	1	1	1	1	1	0.5	1	0.89
4	RC	9	1	1	1	0.5	1	1	1	1	1	1	1	0.95
8	RC	16	1	1	1	0.25	1	1	1	1	0.75	0.25	1	0.84
9	RC	17	0.5	0.5	1	0.25	1	1	1	1	1	0.75	1	0.82
5	RC	21	0	0.75	1	0	0	1	0	0	0.75	1	1	0.50
6	RC	22	0.25	0.25	0.75	0	0.75	0.5	0.25	0	0.25	1	1	0.45
7	RC	25	0	0	1	0	0	0	0	0	0	0	1	0.18
10	RC	30	1	1	1	0	1	1	1	1	1	0	1	0.82
11	RC	37	1	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.59
13	RC	39	0.5	0.75	1	0.5	0.75	1	0.75	0.5	0.75	0.5	1	0.73

Calcul des moyennes en retirant les questions présentant des scores trop divergents par rapport à la moyenne

	E18	E4	E12	E9	E3	E6	E16	E5	E11	E17	E1
Résultat aux questions préalablement testées au format QCM	66.67%	91.67%	94.44%	22.22%	100.00%	100.00%	80.56%	72.22%	72.22%	38.89%	97.22%
Résultat aux questions préalablement testées au format choix ouvert	77.50%	82.50%	90.00%	57.50%	87.50%	92.50%	87.50%	85.00%	87.50%	60.00%	90.00%
Résultat aux questions non testées	58.33%	79.17%	70.83%	8.33%	83.33%	75.00%	58.33%	45.83%	50.00%	29.17%	75.00%
Moyenne QCM	76.01%										
Moyenne choix ouvert	81.59%										
Moyenne non testé	57.58%										