

2023/2024

MASTER 1 / 2

Département MEEF

UE3 Ec 2 Recherche et mémoire

L'EFFET TEST

Les effets d'évaluations régulières sur la qualité des apprentissages des élèves

Présenté par :

Robin GALEZZI

Physique Chimie

Référent mémoire : Mme. DECORSE

DECLARATION DE NON-PLAGIAT

" Je déclare que ce mémoire est le fruit d'un travail personnel et que personne d'autre que moi ne peut s'en approprier tout ou partie.

J'ai connaissance du fait que prétendre être l'auteur de l'écrit de quelqu'un d'autre enfreint les règles liées à la propriété intellectuelle.

Je sais que les propos empruntés à d'autres auteurs doivent figurer entre guillemets. Je m'appuie dans ce mémoire sur des écrits systématiquement référencés selon une bibliographie précise. "

Signature :



Table des matières

INTRODUCTION.....	4
REVUE DE LA LITTERATURE.....	5
Ce qu'il faut en retenir :.....	12
PROBLEMATIQUE	13
PROTOCOLE	14
Le dispositif :.....	14
Etude quantitative :.....	16
Etude qualitative :.....	16
Attentes :.....	18
Analyse des données collectées.....	19
Situation de départ :.....	19
Résultats aux évaluations et aux tests :.....	20
Le stress face aux évaluations :.....	22
Questionnaire de fin :	23
Discussion des résultats :.....	24
Biais expérimentaux :.....	26
Réponse à la problématique :.....	27
CONCLUSION	28
BIBLIOGRAPHIE	29
ANNEXES	31

INTRODUCTION

L'évaluation est souvent associée à une connotation négative, provoquant de la peur et du stress chez les élèves. Cela peut être considéré comme l'une des principales sources de stress pour les étudiants. Lors de mes études à l'Université de Bourgogne, j'ai eu la chance de suivre des cours avec M. THIONNET Alain, un professeur de mécanique. En plus de me donner goût pour sa matière il m'a permis de voir l'évaluation comme une méthode d'apprentissage plutôt que comme une sanction ou un passage obligatoire pour valider un diplôme.

Un nombre croissant d'études se sont intéressées aux avantages et aux effets que les tests peuvent avoir sur la rétention de l'information. Certaines expériences, menées notamment en contexte scolaire, ont conclu que des tests répétés et réguliers sur la matière abordée pouvaient aider à mieux assimiler cette matière et donc à mobiliser les connaissances sur une plus longue durée. On appelle cela l'effet des tests.

Ainsi, la question de recherche qui en découle est la suivante : Quels sont les impacts de la mise en place de tests à chaque séance sur la qualité des apprentissages de physique chimie dans le secondaire ? La présente étude est divisée en deux parties. La première partie de cette étude vise à démontrer l'impact positif des tests en classe sur l'apprentissage des compétences liées à la démarche scientifique, telles qu'elles sont décrites dans les bulletins officiels de l'éducation nationale. En effet, l'utilisation de tests réguliers peut permettre aux élèves de mieux retenir les notions vues en classe sur une durée plus longue, en les aidant à consolider leur compréhension des concepts abordés. Cependant, si les résultats de l'étude ne sont pas concluants, il serait judicieux de recueillir les opinions des élèves afin de comprendre les raisons pour lesquelles le dispositif mis en place n'a pas permis de valider l'hypothèse de l'étude. Dans cette optique, la deuxième partie de cette étude consiste à administrer un questionnaire qualitatif au début et à la fin de l'étude en classe. Ce questionnaire vise à recueillir les opinions de chaque participant de l'échantillon sur leurs ressentis, l'utilité et la pertinence du dispositif mis en place. Enfin, on pourra également se demander si cette expérience aura fait évoluer leur vision de l'évaluation.

REVUE DE LA LITTERATURE

Le phénomène d'effet test, bien qu'il ait été l'objet de nombreuses études par des psychologues expérimentaux, n'est pas largement connu ou apprécié dans l'éducation. En 1620, Bacon a écrit : « Si vous lisez un texte vingt fois, vous ne l'apprendrez pas par cœur aussi facilement que si vous le lisez dix fois en essayant de le réciter de temps à autre et en consultant le texte lorsque votre mémoire fait défaut » (Roediger & Karpicke, 2006).

Classiquement les phénomènes de la mémorisation sont décrits en 3 étapes : l'encodage, le stockage et la récupération. Du phénomène d'encodage dépend la mémorisation à long terme. Quand un apprenant se contente ou n'a le temps que de comprendre, il n'encode pas ou peu il est donc normal qu'il ne se souvienne de rien dans un laps de temps plus ou moins court. Mais il ne suffit pas d'encoder pour se souvenir à long terme. Les modifications des synapses que constituent les phénomènes d'apprentissage s'estompent plus ou moins rapidement si elles ne sont pas consolidées par des sollicitations répétées selon certaines règles. Le mécanisme de la mémorisation peut se résumer comme suit : (Stordeur, 2014)

- Modifications locales dans la synapse / Activation par les kinases, de molécules nommées « facteurs de transcription » / Envoi de signaux vers le noyau / Fixation de ces molécules sur des sites de reconnaissance de l'ADN / Fixation qui déclenche l'expression de certains gènes / Les premiers gènes activés sont dits « gènes précoces » et surtout actifs au niveau des synapses / D'autres gènes dits « effecteurs tardifs » sont aussi activés et envoient des informations vers le noyau,

- **Nécessité d'une sollicitation intense répétée,**

- **Modifications ainsi stabilisées à long terme (mémorisation).**

Une sollicitation importante est nécessaire pour commencer à construire des traces mémorielles du savoir compris (apprentissage). La sollicitation personnelle doit être suffisamment importante pour que les traces se construisent en ayant des chances de rester un certain temps. C'est en évoquant ou sollicitant les traces construites que l'on finit par les rendre définitivement acquises (mémorisation à plus ou moins long terme). Une sollicitation répétée à plus ou moins long terme (évocation personnelle) est indispensable pour consolider les traces en construction (mémoire à long terme). C'est toujours par essai et erreur que l'apprentissage peut se faire. Ce temps d'essais et donc d'erreurs possibles permet les sollicitations suffisantes

à l'ouverture des canaux à calcium assurant les réactions moléculaires nécessaires aux apprentissages (Stordeur, 2014).

La consolidation d'un apprentissage apparaît dès lors que le savoir ou la pratique acquise résonne dans un nouveau contexte. En visualisant et reformulant mentalement comment procéder une prochaine fois, on rend la mémorisation plus durable.

S'exercer permet d'effectuer une action quelconque à la mémoire et la répéter conforte le savoir, créant des raccourcis permettant d'y accéder plus facilement. Aristote a écrit : « s'exercer à se souvenir d'un fait de manière répétée renforce la mémorisation » (Brown et al., 2016). Depuis 1885, les psychologues décrivent une courbe de l'oubli (premier travail publié par Hermann Ebbinghaus en 1885) avec une perte en très peu de temps de ce qui a été lu ou entendu de plus de 70% puis une perte des 30% restant plus lente (Brown et al., 2016). Le défi majeur pour améliorer notre manière d'apprendre serait donc d'entraver ce processus d'oubli. Des données expérimentales montrent que la pratique de la remémoration rend l'apprentissage bien plus efficace qu'une nouvelle exposition aux mêmes faits (Brown et al., 2016). C'est ce que qu'on appelle l'effet test (ou effet d'un travail de remémoration ou de récupération). Cette remémoration pour être optimale doit être **répétée** et **espacée** de manière à ce que le processus de rappel, plutôt que de devenir une récitation par cœur, demande de vrais efforts cognitifs. Des rappels réguliers semblent également aider la mémoire à se consolider. Depuis le début des années 2000, il est donc montré qu'une remémoration répétée peut ancrer si solidement les connaissances et les savoir-faire qu'ils deviennent de véritables réflexes.

Cependant cette méthode reste peu mise en pratique par les professeurs et étudiants des milieux éducatifs traditionnels mais également mal comprise la plupart du temps (enfants, étudiants, parents...).

Une première étude à grande échelle a été publiée en 1917 avec des enfants de CE2, CM2, 6ème et 4ème. Tous les groupes auxquels avait été demandé une récitation silencieuse (60% de leur temps) réussissent mieux à retenir les données que ceux qui s'étaient contentés de lire en continu. (Brown Peter C et al., 2016)

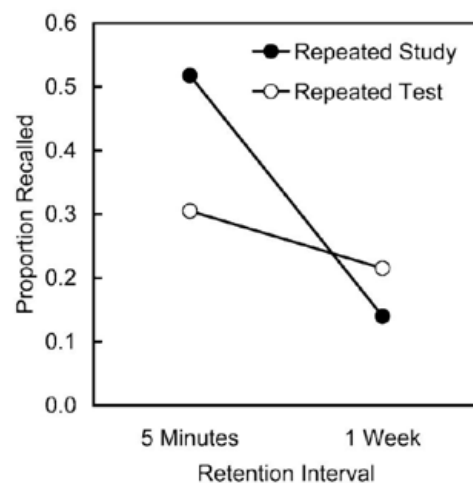
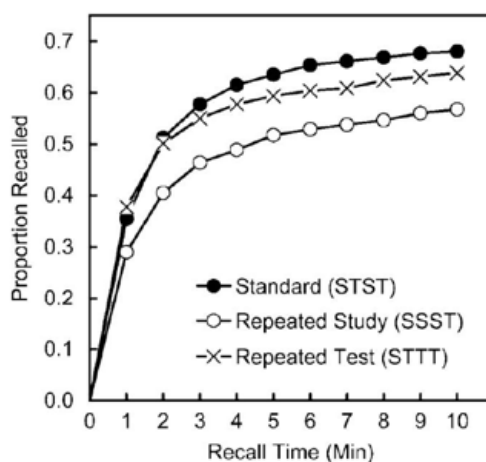
Une deuxième étude en 1939 avec 3000 élèves de 6ème dans l'Iowa a été réalisée amenant deux résultats :

- Plus le premier test était retardé, plus l'oubli était important,

- Une fois qu'un élève avait fait le premier test, l'oubli était arrêté et les résultats des élèves aux évaluations suivantes ne chutaient que faiblement (Brown et al., 2016).

L'intérêt pour l'effet test refit surface en 1967 avec la publication d'une étude montrant que les sujets apprenaient aussi bien en effectuant des tests réguliers qu'en lisant et relisant de manière répétée. Ceci a amené les chercheurs à reconsidérer le potentiel du test comme outil d'apprentissage. En 1978 des chercheurs montrèrent que le travail massé conduisait à de bons résultats si l'évaluation avait lieu dans la foulée de la phase d'apprentissage, mais la perte de mémoire était beaucoup plus importante (50%) lorsqu'un second test avait lieu 2 jours après. Alors que ceux qui avaient effectué des exercices de remémoration n'oubliaient que 13% des données initialement retenues. Lors de recherches ultérieures il en a été conclu que des sessions répétées de pratiques de récupération sont généralement préférables à un entraînement unique, en particulier si ces entraînements sont espacés dans le temps (Brown et al., 2016).

Un projet d'étude a débuté en 2005 sur le terrain dans un collège de Columbia dans l'Illinois sur 3 semestres (1 an et demi). Roediger & Karpicke (2006) ont mené une première expérience avec les trois conditions de Tulving (étude standard, études répétées et tests répétés), en utilisant 40 mots et un taux de présentation de 3s. Ainsi, ils ont reproduit le résultat de base de Tulving réalisé en 1967 selon lequel les courbes d'apprentissage pour ces trois conditions sont remarquablement similaires, bien qu'ils aient trouvé une différence favorisant la condition standard. Ainsi, malgré le fait que les sujets dans la condition d'étude répétée avaient étudié la liste 15 fois, 1 semaine plus tôt et ceux dans la condition d'essai répété l'avaient étudiée seulement 5 fois, le rappel retardé était plus grand pour ce dernier groupe. Ce résultat montre le pouvoir des tests pour améliorer la rétention à long terme.

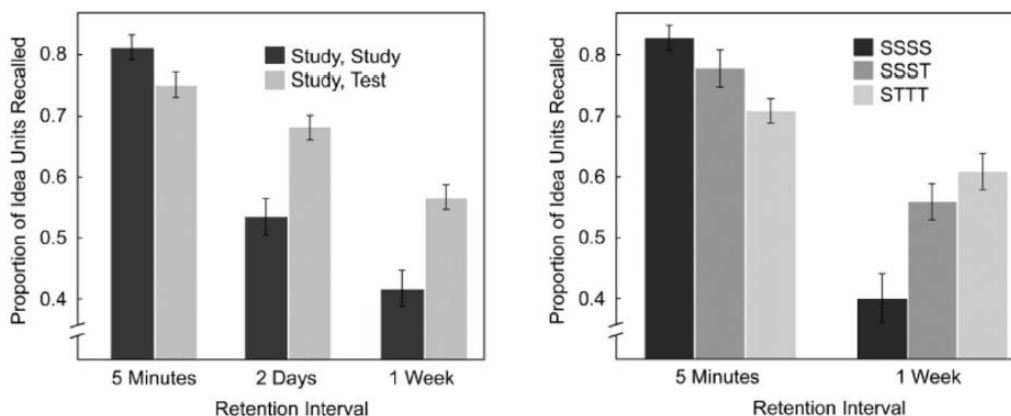


Gauche : Rappel cumulatif d'un test de rétention final effectué une semaine après l'apprentissage initial. Les étiquettes de raccourci indiquent l'ordre des périodes d'étude (S) et d'essai (T).

Droite : Proportion de mots rappelés immédiatement (5 min) et retardés (7 jours) tests de rétention après des études répétées ou des tests répétés. Les données sont estimées par Wheeler, Ewers et Buonanno (2003).

Pour résumer : L'étude répétée est bénéfique pour les tests donnés peu après l'apprentissage, mais sur les tests retardés avec des intervalles de rétention mesurés en jours ou en semaines, les tests antérieurs peuvent produire des performances plus élevées que les études antérieures. Les tests réduisent l'oubli du matériel récemment étudié, et les essais multiples ont un plus grand effet pour ralentir l'oubli qu'un seul essai.

Une deuxième expérience a étudié les effets des études répétées et des tests répétés sur la rétention ultérieure. Les sujets ont étudié des passages comme suit : SSSS, SSST, STTT. (S : Etude ; T : Test) Ils ont passé un test de rappel final cinq minutes ou une semaine après cette séance d'apprentissage. Les résultats sont montrés dans la figure de droite. Le groupe SSSS est celui qui a le plus tendance à l'oubli (52 %), suivi par le groupe SSST (28 %), et le groupe des tests répétés (STTT) a le moins tendance à l'oubli (10 %) sur 1 semaine.



Gauche : Proportion moyenne d'unités d'idées rappelées après un intervalle de rétention de 5 minutes, 2 jours ou 1 semaine. Les barres d'erreur représentent les erreurs types des moyennes.

Droite : Proportion moyenne d'unités d'idées rappelées lors d'un test final 5 min ou 1 semaine après l'apprentissage. Les étiquettes de raccourci indiquent l'ordre des périodes d'étude (S) et d'essai (T).

Pour résumer : Les résultats sont bien supérieurs sur les notions ayant fait l'objet de questionnaires par rapport à celles qui n'avaient fait l'objet que de révision factuelle. Les résultats de ces dernières n'étaient pas significativement meilleurs que les notions non revues en classe.

En 2007, l'étude fut étendue à un cours de sciences en 4ème. Les résultats indiquent que l'effet bénéfique des tests dure au moins 9 mois. Les tests ont augmenté le rendement des élèves aux examens de l'unité, de 79 % à des niveaux de plus de 90 %. Le gain de performance lié au questionnaire représentait un changement d'une note de C + à une note de A- sur la distribution de notes typiques à cette école (McDaniel et al., 2011) et (Brown et al., 2016).

Les tests pratiqués n'ont pas été vécus comme des moments pénibles par les élèves, 64% ont déclaré que ces tests diminuaient leur anxiété devant le devoir de fin de chapitre et 89% avaient eu le sentiment que cela facilitait leur apprentissage. Le principal du collège dans lequel a été réalisée l'étude, M. Roger Chamberlain déclara « la pratique de la récupération a un impact significatif sur l'apprentissage des enfants. Cela nous prouve bien que c'est un outil précieux et que les professeurs seraient bien avisés de l'introduire dans leur pratique pédagogique » (Brown et al., 2016).

Ces deux articles couplés (McDaniel et al., 2011) et (Roediger et al., 2011) furent les premiers à présenter des résultats d'expériences rigoureuses sur les bénéfices des tests pour les résultats obtenus par des élèves de collège en sciences et en sciences humaines et sociales. Les résultats montrèrent que les tests produisaient une amélioration sensible par rapport à l'absence de test ou à la simple révision, à la fois sur les évaluations ponctuelles, trimestrielles et annuelles.

Le report du premier essai améliore la rétention à long terme, peu importe la façon dont les essais répétés sont espacés. Le facteur important pour la promotion de la rétention à long terme est de retarder la récupération initiale pour la rendre plus difficile. L'allongement de l'intervalle entre les essais répétés a eu peu d'effet sur la rétention à long terme (Karpicke & Roediger, 2007). Par conséquent il n'est pas préférable de tester les élèves immédiatement à l'issue de la leçon mais d'attendre la séance suivante.

Dans un sondage auprès de 170 étudiants d'un établissement privé, 84 % des étudiants ont déclaré utiliser la lecture répétée comme méthode d'étude, et 55 % ont indiqué la relecture comme leur stratégie numéro une. Une série de 4 expériences montre clairement que lorsque l'on utilise des matériaux pertinents sur le plan éducatif et des évaluations sommatives, la

relecture immédiate n'est pas une méthode d'étude efficace. Il y avait une absence constante d'effets de relire lors des quatre expériences. Cette absence constante d'effets de la relecture collective montre que la relecture immédiate n'améliore pas de façon fiable le rendement des évaluations sommatives habituellement utilisées en classe (Callender & McDaniel, 2009).

L'un des effets indirects des tests est l'effet induit de la pratique de récupération par encodage ultérieur. Tenter de récupérer des éléments peut améliorer l'encodage ultérieur de ces éléments même lorsque la tentative de récupération échoue et que la rétroaction n'est pas donnée. Cet effet, connu sous le nom d'apprentissage potentialisé par test, a d'abord été identifié par Izawa en 1966, qui a suggéré que plus de tentatives de récupération conduisent à une plus grande potentialisation. Faire une interrogation même lorsqu'on ne parvient pas à donner les bonnes réponses améliore l'acquisition des savoirs lors d'un nouvel épisode d'étude (Arnold & McDermott, 2013).

Après une interrogation, les étudiants passent plus de temps à réétudier les points qu'ils ont manqués et apprennent mieux que leurs camarades qui retravaillent la matière sans avoir été interrogés. Ceux qui mettent l'accent sur la relecture mais pas sur l'auto-interrogation ont tendance à surestimer leur maîtrise. Les étudiants qui ont été interrogés possèdent donc un double avantage :

- Un sens plus aigu de ce qu'ils savent et ne savent pas,
- Un renforcement de leurs apprentissages.

Bien sûr, il peut y avoir des effets négatifs des tests, comme l'anxiété liée à ceux-ci. Nous soupçonnons que si les tests en classe étaient plus routiniers et qu'il y avait peu de « gros tests » qui comptaient pour la majeure partie de la note d'un élève, des phénomènes comme l'anxiété liée aux tests diminuerait par l'accoutumance (Roediger & Karpicke, 2006).

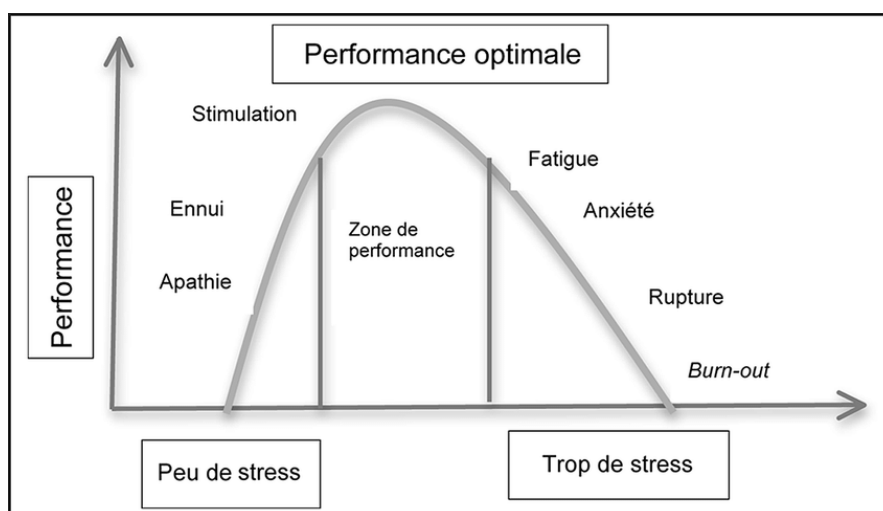
Everson et al. (1993), cités par Viau (1995), ont établi un classement des matières scolaires en fonction du niveau d'anxiété déclaré par les élèves. Cette étude a démontré que les matières scientifiques suscitaient plus d'anxiété chez les élèves. Sur cette liste, les sciences physiques apparaissent comme la discipline la plus anxiogène, suivies par les mathématiques.

Le sentiment d'efficacité personnelle concerne les croyances individuelles quant à la capacité d'accomplir des performances spécifiques (Bandura, 2003). Il influence les choix d'activités, l'investissement dans la réalisation des objectifs, la persistance des efforts et les réponses émotionnelles face aux obstacles. La réussite dans des expériences concrètes renforce

cette croyance, tandis que les échecs la diminuent, en particulier si la tâche n'est pas trop facile. De plus, les états émotionnels, tels que l'anxiété associée à une faible performance, peuvent engendrer des doutes quant aux compétences personnelles, conduisant éventuellement à l'échec. En revanche, des états émotionnels positifs favorisent la confiance en soi et la réussite. Je m'interroge et espère que la réussite lors de petites évaluations régulières pourra stimuler le sentiment d'efficacité personnelle des élèves, encourageant ainsi leur engagement accru dans la matière.

Parmi les impacts cognitifs, notons qu'un haut niveau d'anxiété aux évaluations entraverait l'encodage, l'organisation et la mémorisation des informations. Au fil du temps, les élèves qui souffrent d'anxiété aux évaluations auront tendance à attribuer leurs échecs à des facteurs intrinsèques (peu de mémoire, faible niveau d'intelligence, etc.) ou extrinsèques (évaluation trop difficile, peu de soutien dans les apprentissages, questions jugées injustes, etc.) hors de leur contrôle (Cassady, 2004), ce qui laisse entrevoir des impacts sur la motivation ou l'estime de soi.

Par ailleurs, le lien entre le stress et la performance a été établi il y a plus d'un siècle par la loi de Yerkes-Dodson (1908). Celle-ci est représentée sous la forme d'une courbe en « U » inversé, la relation entre le niveau d'éveil et la performance cognitive. Cette théorie indique l'importance d'un niveau minimal de stress pour atteindre le seuil de performance optimale. Au contraire, lorsque le niveau de stress est trop élevé, c'est la performance cognitive qui se voit être amoindrie.



Selon Viau (1995), la distinction entre l'anxiété et le stress est difficile à établir. Certains chercheurs les considèrent comme synonymes. Cependant, les études portant sur le stress se sont davantage intéressées aux différents éléments qui perturbent l'environnement (passation

d'un examen, prise de parole en classe) tandis que les études basées sur l'anxiété se sont penchées davantage sur les facteurs cognitifs et émotionnels. C'est la raison pour laquelle j'ai décidé de parler du stress plutôt que de l'anxiété par la suite. Il est pertinent de préciser que ce stress n'affecte pas quotidiennement et avec la même intensité tous les élèves. Certaines situations vont paraître plus stressante que d'autres. Par exemple, une évaluation formative favorisera un stress plus faible qu'un examen de fin d'année.

Ce qu'il faut en retenir :

Le fait de se remémorer une nouvelle connaissance ou une nouvelle compétence constitue un outil puissant pour un apprentissage et une mémorisation à long terme. Une remémoration qui demande des efforts construit un apprentissage et une mémorisation à long terme plus solides : lorsque le cerveau doit travailler dur, ce qu'on apprend reste mieux ancré. Le test est un outil puissant pour améliorer l'apprentissage. De nombreuses études en laboratoire ont démontré ce point, et les quelques applications systématiques en classe ont réussi à améliorer les performances. Si les enseignants déterminent les connaissances et les compétences essentielles qu'ils veulent que leurs élèves connaissent après avoir quitté la classe, ces points peuvent être soulignés en classe et mis à l'essai à intervalles réguliers pour s'assurer que les élèves acquièrent ces connaissances. Les tests fréquents ont non seulement un effet direct sur l'apprentissage, mais ils devraient aussi encourager les élèves à étudier davantage, à s'engager continuellement dans la matière, à éprouver moins de stress et probablement même à obtenir de meilleurs résultats aux tests normalisés. Une pratique de remémoration répétée rend non seulement les éléments mémorisés plus durables mais produit aussi un savoir qui peut être mobilisé plus rapidement, dans des contextes plus divers et ainsi appliqué à travers une gamme plus vaste de problèmes. Les apprenants qui font des interrogations saisissent mieux leur niveau que ceux qui se contentent de relire la matière et améliorent l'acquisition des savoirs lors d'un nouvel épisode d'étude. De plus, elles permettent à l'enseignant de repérer les lacunes et erreurs de compréhension et d'adapter la suite du cours pour corriger celles-ci. Les étudiants qui ont régulièrement des interrogations à faible enjeu finissent par intégrer cette pratique à leur propre méthode de travail.

PROBLEMATIQUE

Les recherches sur l'effet test ont profondément influencé ma perception de ce que de simples petites évaluations en début de cours ont pu m'apporter au cours de ma scolarité et peuvent apporter aux élèves. Ainsi ma question initiale sur les effets de la mise en place de tests à chaque séance sur la qualité des apprentissages en physique chimie au niveau secondaire. Je réalise maintenant que ces évaluations ne sont pas seulement un moyen de pousser les élèves à étudier régulièrement, mais qu'elles constituent également un outil puissant pour favoriser les apprentissages. Cela me conduit à me poser la question suivante de manière plus approfondie :

Les évaluations brèves et à faible enjeu en début de séance peuvent-elles améliorer les apprentissages des élèves et réduire leur stress lié aux évaluations ?

Les études sur l'effet test ont montré que les évaluations régulières, même lorsqu'elles sont de courte durée et sans conséquences majeures, ont un impact significatif sur la rétention des connaissances et la consolidation des apprentissages. En réalisant des évaluations fréquentes au début de chaque séance, les élèves sont incités à récupérer et à activer les informations préalablement apprises, ce qui renforce leur mémorisation. De plus, ces évaluations régulières me permettront de repérer rapidement les lacunes et les difficultés des élèves, ce qui me donnera l'occasion d'adapter mon enseignement. En tant qu'enseignant débutant ces petites évaluations pourront m'être très bénéfiques pour corriger rapidement les lacunes de mon enseignement.

PROTOCOLE

Dans le but de tester l'effet de la récupération sur les apprentissages, il sera mis en place un rituel de récupération en début de chaque cours. En effet, des études ont montré que la récupération est un moyen efficace de consolider les connaissances et de les rendre plus durables. Dans un premier temps nous verrons le dispositif qui sera mis en place puis les études quantitatives et qualitatives que l'on pourra faire autour de ce dispositif, enfin seront abordés les biais expérimentaux et les attentes propres de cette étude.

Le dispositif :

J'ai conduit mon expérience au sein d'une classe de première en enseignement scientifique, j'ai la chance d'avoir cette classe divisée en deux groupes, chacun composé de 16 élèves. L'un de ces groupes a été désigné comme le groupe test, tandis que l'autre fonctionne en tant que groupe témoin. Cette approche me permet d'évaluer de manière comparative les résultats et les effets du dispositif mis en place dans le groupe test.

Il est demandé aux élèves en début de cours de ne sortir que le nécessaire à l'évaluation, pour leur donner un signal clair que le temps est consacré à l'évaluation et à la récupération, plutôt qu'à la préparation pour le cours. Une courte évaluation portant sur une notion importante du programme de l'année en cours et déjà étudiée en classe, est posée sur leur table. Cela permettra aux élèves de se concentrer sur des connaissances précises et pertinentes, tout en les encourageant à réviser régulièrement pour ne pas être pris au dépourvu. (Voir les exercices proposés en annexe)

Il est également important de noter que les évaluations ne doivent pas être toutes sous la forme de QCM. Bien que les QCM soient utiles pour évaluer la compréhension globale d'un sujet, ils ne sont pas aussi efficaces que les interrogations qui demandent à l'élève de fournir une réponse avec des phrases rédigées, ou avec un schéma. Cela demande un effort cognitif plus important de la part de l'élève, et favorise donc une meilleure consolidation de ses connaissances. Je suis conscient que la correction d'un nombre important de copies peut être chronophage, c'est pourquoi j'ai décidé de m'autoriser l'utilisation occasionnelle de questionnaires à choix multiples.

Les créneaux horaires de la classe constituant le groupe test sont situés, juste après une récréation ou alors je suis déjà dans la même salle lors de mon cours précédent avec le groupe témoin. Cette planification me permet d'avoir suffisamment de temps pour préparer la salle de

classe et les outils nécessaires pour les tests. De plus, cela me permet de m'assurer que tout est en ordre avant l'arrivée des élèves, minimisant ainsi les perturbations sur le temps de classe. En somme, ce type d'aménagement horaire me permettrait d'être plus efficace dans la gestion de ma classe.

Il est important de souligner que la récupération doit être régulière pour être efficace. Le fait de pratiquer ce rituel en début de chaque cours permet aux élèves de réviser régulièrement et d'assimiler les connaissances petit à petit, plutôt que de tout apprendre en une seule fois avant un examen.

Le groupe d'élèves formant le groupe test se caractérise par son calme et sa coopération, ce qui a grandement facilité la mise en place du rituel. Cette atmosphère favorable a permis de minimiser les pertes de temps pendant la période de cours, rendant l'instauration du rituel plus efficace et bénéfique pour l'expérience globale.

Il est primordial que les évaluations soient corrigées rapidement et annotées afin que les élèves puissent bénéficier au maximum de l'expérience. Ainsi, des séances de correction pour ces évaluations sont organisées juste à l'issue du test et les copies rendues dès la séance suivante. Ces séances de correction permettent aux élèves de comprendre leurs erreurs et d'identifier les domaines dans lesquels ils doivent se concentrer pour améliorer leur compréhension et leurs compétences. En outre, ces séances et la correction des copies m'offriront également une occasion d'identifier les lacunes dans l'enseignement et d'ajuster mes cours pour mieux répondre aux besoins des élèves avant l'évaluation sommative.

Le contenu didactique retenu pour l'expérimentation est le suivant :

- Appliquer la loi de Wien pour déterminer la température de surface d'un objet assimilé à un corps noir à partir de la longueur d'onde d'émission maximale.
- La puissance radiative reçue du Soleil par une surface plane est proportionnelle à l'aire de la surface et dépend de l'angle entre la normale à la surface et la direction du Soleil.
- Calculer la longueur du méridien terrestre par la méthode d'Ératosthène (Proportionnalité).

Seules des notions importantes et bloquantes pour les élèves font l'objet d'évaluations systématiques. Il est crucial, autant que cela est possible étant donné le très court délai de

l'expérimentation, que chaque notion soit testée à plusieurs reprises pour consolider les traces mémorielles. C'est pourquoi je fais le choix de me concentrer seulement sur quelques notions qui me semblent plus importantes et englobant un panel large de compétences, que sur un panel trop large de notions. Les notions dites « importantes » pourront être évaluées sur l'ensemble de l'année à partir du moment où elles ont été vues en classe et pas seulement lors de la séquence portant sur cette même notion. Plus l'année scolaire avance plus le panel de notions évaluables lors des tests de début de séance augmente. Il est important de noter que le choix des notions importantes du programme est basé sur une analyse du BO de première enseignement scientifique et des difficultés des élèves observées en classe. Cette analyse permet de déterminer les connaissances qui me semblent clés qui doivent être comprises et maîtrisées par les élèves pour assurer leur réussite.

Etude quantitative :

L'objectif principal de cette étude est d'évaluer si l'utilisation de l'effet test permet une meilleure rétention des notions de physique chimie dans l'enseignement secondaire. L'étude se compose de deux parties, une quantitative et une qualitative. La première partie vise à comparer les résultats aux évaluations sommatives communes d'une classe sans le dispositif de test (qu'on appelle groupe témoin) et d'une classe équivalente bénéficiant de ce dispositif (qu'on nomme groupe test) pour mesurer l'impact des tests sur la rétention des notions qui auront été testées. Idéalement, les deux groupes auraient dû présenter un niveau global similaire, cependant, ce n'est pas le cas. Nous examinerons dans l'analyse des données la manière dont nous avons remédié à ce problème. Je n'ai pas la possibilité de sélectionner mon échantillon afin que mon groupe témoin ait le même niveau de départ que mon groupe test cela engendre un biais qui ne sera pas négligeable lors de l'analyse des résultats.

Etude qualitative :

La seconde partie consiste à interroger les élèves sur leur perception de cette expérience pour recueillir leurs avis et ressentis. Ainsi, la première partie permet de vérifier les hypothèses émises, tandis que la seconde partie permet de comprendre comment le dispositif est perçu par les sujets. Les élèves seront sollicités pour remplir deux questionnaires (voir questionnaire 1 et questionnaire 2 en annexe). Le premier a été complété avant la mise en place de l'expérience, tandis que le second a été rempli après quelques mois d'expérimentation. Ces questionnaires sont basés sur l'échelle de Likert et ont pour objectif d'évaluer les sentiments des élèves

concernant certains aspects précis, ainsi que de déterminer si leur point de vue a évolué suite à l'expérimentation.

L'échelle de Likert consiste en une série d'énoncés auxquels les répondants doivent indiquer leur degré d'accord ou de désaccord. J'ai utilisé une échelle à cinq points allant de "tout à fait en désaccord" à "tout à fait d'accord". Les résultats obtenus à partir de l'échelle de Likert permettent de quantifier les opinions ou attitudes des élèves.

Plus précisément, les questionnaires se sont concentrés sur les éléments suivants :

- Les méthodes de travail des élèves (telles que la lecture, l'auto-évaluation, les fiches de révision, etc.) : il a été observé que seulement 11% des élèves ont recours aux tests pour apprendre. De plus, ceux qui se testent déclarent principalement le faire pour identifier les points sur lesquels ils ne sont pas sûrs afin de se concentrer sur ces aspects lors de leurs révisions. Peu d'élèves réalisent que le simple fait de se souvenir activement d'une information favorise sa mémorisation (Brown et al., 2016). Lors d'un sondage mené auprès de 170 étudiants d'un établissement privé, 84% des participants ont déclaré utiliser la lecture répétée comme méthode d'étude, tandis que 55% ont indiqué que la relecture était leur stratégie de révision numéro une (Callender & McDaniel, 2009).
- La perception des élèves quant à la compréhension et l'apprentissage d'un concept : lorsqu'un apprenant se contente simplement de comprendre un concept sans l'encoder de manière suffisante, il est normal qu'il ne parvienne pas à se rappeler de l'information sur le long terme (Stordeur, 2014).
- Le stress des élèves face à l'évaluation formative et sommative.

Il est important que ces évaluations systématiques ne soient pas vécues comme des moments pénibles et anxiogènes. Il est important que le concept de l'effet test soit expliqué aux élèves participant à l'étude. Il leur a également été expliqué que ces évaluations n'auront que peu d'impact sur les résultats trimestriels. Roediger & Karpicke (2006) soupçonnent que si les tests en classe étaient plus routiniers et qu'il y avait que peu de « gros tests » qui comptaient pour la majeure partie de la note d'un élève, des phénomènes comme l'anxiété liée aux tests diminuerait par l'accoutumance. Ces propos sont confirmés dans Brown et al. (2016) qui nous dit que des tests pratiqués en 2007 dans le même collège que celui de l'étude de 2006 n'ont pas été vécus comme des moments pénibles par les élèves : 64% ont déclaré que ces tests diminuaient leur anxiété devant le devoir de fin de chapitre et 89% avaient eu le sentiment que

cela facilitait leur apprentissage. Dans l'optique de vérifier ces propos une question sur le niveau de stress précède chaque évaluation formative et sommative. Cette question « Je suis stressé par cette évaluation » sera basée sur une échelle de Likert allant de "Non pas du tout" à "Oui beaucoup".

Chaque évaluation sera notée sur 5 points. Il revêt une importance particulière d'accorder aux élèves le droit à l'erreur, et il est essentiel qu'ils en soient conscients afin de réduire le stress lié à l'évaluation. C'est dans cette optique que seules les quatre meilleures performances seront prises en compte pour établir une note sur 20, excluant ainsi les résultats moins favorables de la note trimestrielle. Cette approche vise à encourager l'apprentissage tout en offrant une marge de manœuvre aux élèves pour surmonter les difficultés et progresser.

Un rituel ?

Il me semble important de ritualiser ces tests. Les rituels ont un impact positif sur le sentiment de sécurité de l'élève et offrent un cadre structurant et ordonné pour se repérer. Les aspects répétitifs et prévisibles des rituels aident les élèves à anticiper la durée d'un certain travail et à prévoir ce qui va se passer. De plus, cette maîtrise du "connu" permet à l'élève de développer une confiance en soi qui sera très utile pour acquérir de nouvelles connaissances (Jeffrey, 2015). Les rituels peuvent aider à gérer le temps de manière efficace afin que ces évaluations systématiques n'empiètent pas de manière trop importante sur le temps de cours.

Attentes :

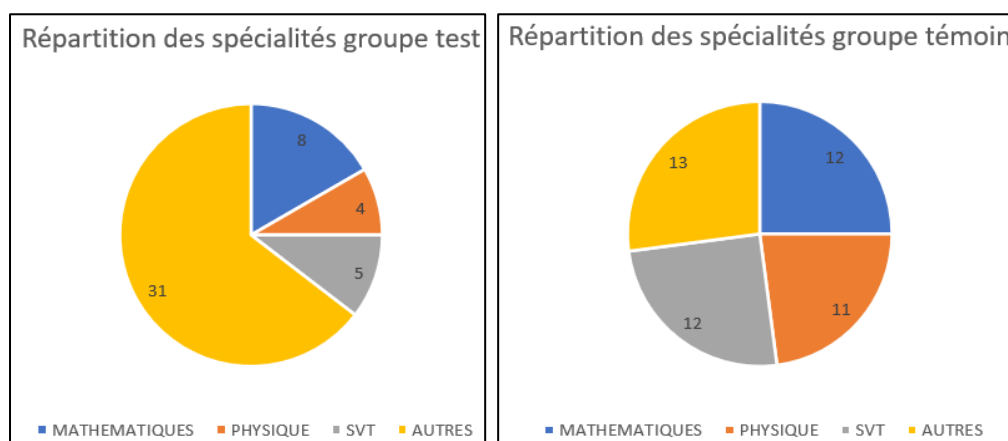
On peut s'attendre à ce que cette expérience apporte plusieurs avantages. Tout d'abord, qu'elle permette une meilleure rétention des notions les plus importantes du programme, ce qui aurait un impact positif sur la réussite des élèves. De plus, en proposant une évaluation régulière et en la présentant comme un outil d'apprentissage, j'espère encourager les élèves à l'utiliser dans leur méthode d'apprentissage personnelle et à prendre conscience de son importance pour progresser. En outre, j'aimerais que cette expérience aide à réduire le stress vis-à-vis de l'évaluation sommative et de l'évaluation en général, en montrant aux élèves qu'il s'agit d'un processus d'apprentissage continu plutôt que d'une simple évaluation de leurs connaissances. Je pense que cela peut également encourager les élèves à étudier de manière plus continue, plutôt que de manière condensée juste avant une évaluation importante.

Analyse des données collectées

Situation de départ :

Il est important de noter qu'un cas particulier s'est présenté dans l'étude du groupe témoin, les résultats d'une élève n'ont pas été pris en compte en raison de son Plan d'Accompagnement Personnalisé (PAP) pour de l'anxiété. Cette situation est notable car elle introduit une variable externe importante dans une étude portant sur le stress. Étant donné que cette élève était incluse dans le groupe témoin, il est impossible d'étudier l'impact du dispositif sur sa situation. De plus, un élève a été retiré du groupe test en raison d'un absentéisme très élevé, ce qui a compromis sa participation effective à l'expérience. Par conséquent, chaque groupe est composé de 15 élèves.

Il est intéressant de noter un déséquilibre significatif dans le choix des spécialités entre les deux groupes. Le groupe témoin semble être davantage orienté vers les spécialités scientifiques, tandis que le groupe test se tourne vers d'autres spécialités. Ce déséquilibre peut potentiellement avoir un impact sur les résultats de l'expérience, étant donné que l'expérience se déroule en première enseignement scientifique. On peut s'attendre à ce que le groupe témoin, avec une plus grande proportion d'élèves orientés vers les spécialités scientifiques, ait un avantage dans la compréhension des concepts scientifiques enseignés.

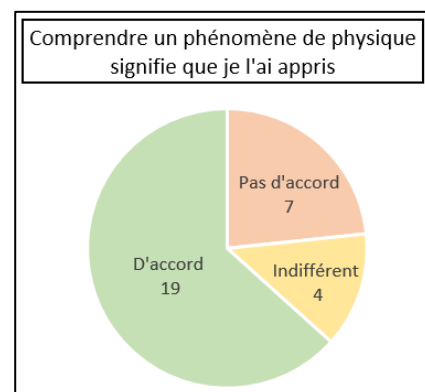
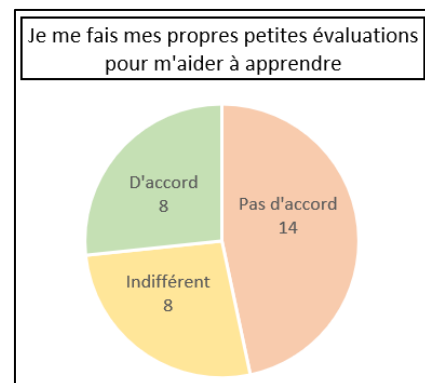


L'expérience a débuté avec la première évaluation de l'année. Un groupe s'est nettement démarqué de l'autre, avec un écart de 3,4 points entre les moyennes de chacun des groupes, cela confirme les attentes. Dans le but de mieux comprendre l'effet test et d'explorer la possibilité de réduire cet écart, il a été choisi de désigner le groupe ayant obtenu les résultats les plus faibles comme groupe test. En optant pour cette approche, l'espoir était de favoriser une progression

au sein du groupe test, tout en évaluant l'impact de cette intervention sur les résultats académiques.

À la suite de la réalisation du premier devoir, une explication détaillée de l'expérience a été présentée à chaque groupe. Par la suite, le questionnaire 1 leur a été distribué. Ce questionnaire visait à recueillir des informations sur :

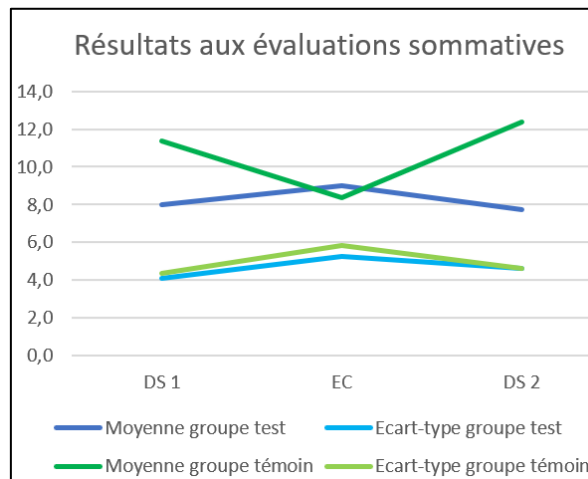
- **Les méthodes d'apprentissage** : Les résultats indiquent que les élèves du groupe test semblent favoriser la relecture des cours comme méthode d'apprentissage, tandis que ceux du groupe témoin ont une préférence pour l'utilisation de fiches. Cependant, il est important de noter que seulement 8 élèves sur 30 adoptent une méthode d'apprentissage par auto-évaluation, alors qu'il est connu de la littérature que c'est une stratégie d'apprentissage plus efficace que la relecture ou les fiches de révision.
- **La compréhension des concepts** : La moitié des élèves associent la compréhension d'un phénomène de physique à l'apprentissage. Cependant, il est important de noter que comprendre un concept ne garantit pas nécessairement la capacité à l'expliquer ou à l'appliquer lors d'un exercice.
- **Le niveau de stress** : Les résultats montrent que la majorité des élèves des deux groupes se sentent stressés à l'approche d'une évaluation notée. En revanche, une majorité d'élèves des deux groupes se déclarent non stressés à l'approche d'une évaluation non notée.



Résultats aux évaluations et aux tests :

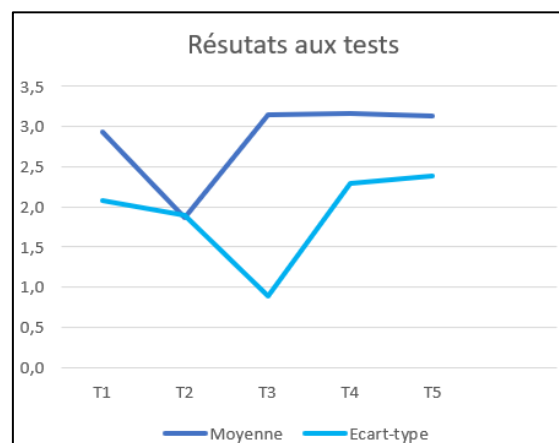
Pour la première partie de l'étude, j'ai comparé les résultats obtenus par le groupe test et le groupe témoin avant la mise en place du dispositif et après quelques semaines de rituel de test avec le groupe test pour vérifier si l'hypothèse selon laquelle l'utilisation de l'effet test dans l'apprentissage des notions de physique chimie correspond aux résultats obtenus par Roediger et al. (2011) ; à savoir que le groupe test produirait une meilleure rétention des informations. Le délai imparti pour mener l'expérience était limité, ce qui a permis de réaliser seulement quatre tests avant de passer à l'évaluation sommative suivante. Cette évaluation était une

épreuve commune à toutes les classes de première enseignement scientifique du lycée, elle englobait l'ensemble des notions abordées depuis le début de l'année scolaire. Ainsi, cette évaluation représentait un moment crucial pour mesurer la progression des élèves et évaluer l'efficacité de l'intervention menée sur le groupe test. En raison de sa portée globale et de son importance dans l'évaluation des acquis des élèves, cette épreuve constituait un indicateur significatif pour analyser l'impact de l'expérience sur les performances académiques.



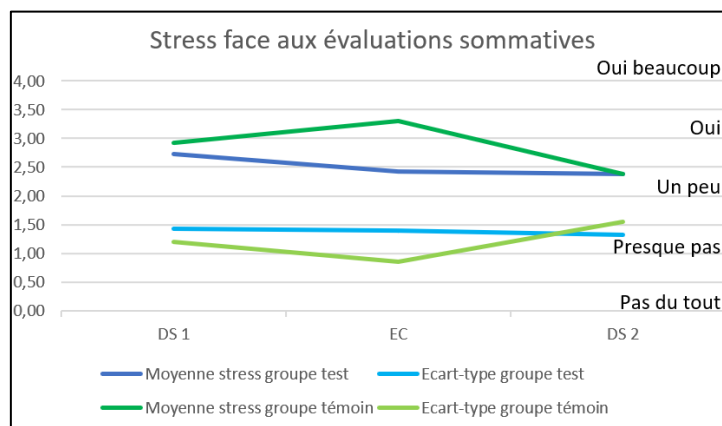
Il est intéressant de constater que les premiers résultats de l'étude ont montré une inversion des performances entre le groupe test et le groupe témoin dès la deuxième évaluation sommative, à savoir l'épreuve commune (EC). Cette observation semble indiquer que l'intervention mise en place auprès du groupe test a eu un impact initial positif sur leurs performances académiques, ce qui est encourageant.

Cependant, une nouvelle inversion des performances a été observée quelques semaines plus tard lors d'une autre évaluation sommative (DS 2), avec le groupe test affichant des résultats nettement inférieurs par rapport au groupe témoin lors de cette évaluation. Cette évolution inattendue soulève des questions.

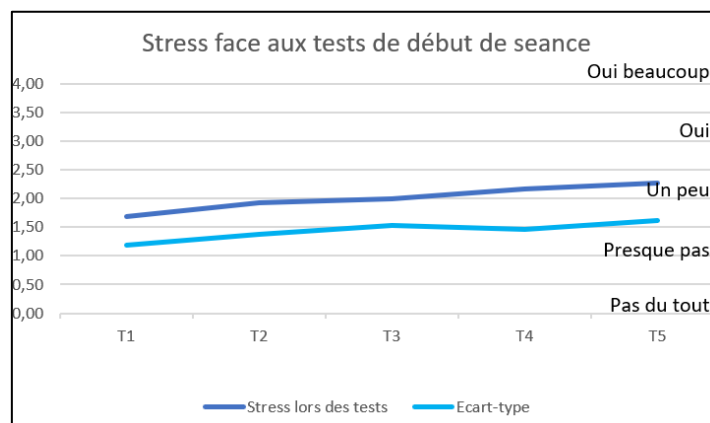


Mis à part le test 3 qui était un QCM, on remarque une forte hétérogénéité dans les résultats lors des tests de chaque séance, ce qui était attendu étant donné que chaque test portait sur une seule notion ou un seul exercice. Cette diversité dans les performances des élèves peut s'expliquer par le fait qu'il n'était pas possible de se rattraper sur d'autres exercices lors de ces évaluations, contrairement à des épreuves plus complexes ou diversifiées. Malheureusement, certains élèves ne semblent pas accorder d'importance à ces évaluations, comme en témoignent le fait que 4 élèves sur 15 ne se soient pas prêtés sérieusement à ces tests en témoigne de nombreuses copies rendues blanches lors de ces évaluations. Ces comportements sont associés à un niveau de stress indiqué comme "pas du tout" ou "presque pas". Cela suggère que d'autres facteurs, tels que le désintérêt, pourraient être en jeu.

Le stress face aux évaluations :



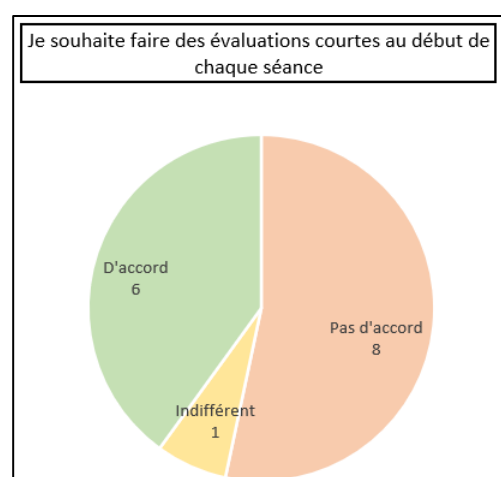
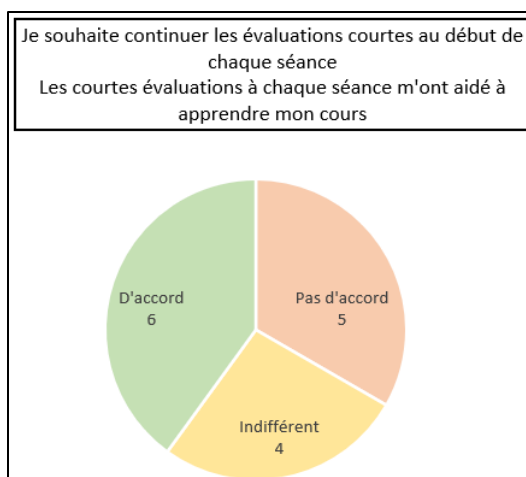
Le niveau de stress lié aux évaluations sommatives a montré des tendances différentes entre le groupe test et le groupe témoin. Pour le groupe test, le niveau de stress a diminué au fil du temps. En revanche, le niveau de stress du groupe témoin est apparu plus variable, avec un pic important observé chez la majorité des élèves lors de l'épreuve commune.



Le niveau de stress des élèves face aux tests de début de séance a légèrement augmenté au fil du temps. Pour le groupe test, on observe que ce niveau de stress est pratiquement identique entre le dernier test et la dernière évaluation sommative. Il est intéressant de souligner que certains élèves du groupe test continuent à indiquer un niveau de stress élevé ou très élevé, même lors du cinquième test, malgré leurs résultats aux quatre tests précédents leurs garantissaient déjà de très bonnes notes (de 17 à 20/20). Ces observations semblent contradictoires avec les réponses aux questionnaires 1, où seulement un élève de ce groupe se dit stressé face à une évaluation non notée. Cette contradiction suggère que le niveau de stress des élèves ne varie guère selon le contexte de l'évaluation.

Questionnaire de fin :

Après quelques semaines d'expérimentation, un second questionnaire leur a été transmis, reprenant les mêmes questions que le premier, avec deux questions supplémentaires pour le groupe test sur leur sentiment vis-à-vis des tests en début de séance et leur volonté de continuer ou non. Pour le groupe témoin, une question a été ajoutée pour savoir s'ils souhaitaient avoir des tests en début de chaque séance. Après analyse des réponses fournies par les deux groupes, on observe très peu de changement de comportement vis-à-vis de leurs méthodes d'apprentissage. Les tests en début de séance ne semblent pas avoir favorisé l'adoption de l'auto-évaluation dans leurs méthodes d'apprentissage. Cependant, si l'on ne prend pas en compte les élèves qui ne se prononcent pas sur la question, une petite majorité d'élèves du groupe test s'accorde pour dire que les tests en début de séance leur ont été bénéfiques dans l'apprentissage de leurs leçons et qu'ils souhaitent continuer ces moments d'évaluation en début de séance. Dans le groupe témoin, lorsqu'on leur a posé la question de savoir s'ils souhaitaient effectuer de petites évaluations en début de séance, la majorité des élèves étaient en désaccord.



Discussion des résultats :

Effectivement, les résultats du questionnaire de départ de l'expérience révèlent que la majorité des élèves n'adoptent pas les méthodes d'apprentissage reconnues comme étant les plus efficaces par la littérature, comme l'auto-évaluation. De plus, il est également observé que l'apprentissage demeure souvent superficiel, avec une majorité d'élèves se contentant de comprendre un phénomène ou un exercice de physique, sans nécessairement vérifier s'ils sont capables de le refaire seuls ou de l'expliquer. Cette tendance suggère que les élèves peuvent manquer de profondeur dans leur compréhension des concepts et ne pas développer leurs compétences en résolution de problèmes.

Il est décevant de constater que la mise en place du dispositif n'a pas permis d'encourager de meilleures pratiques lors des apprentissages. Un travail plus axé sur l'auto-évaluation, en proposant des méthodes et des ressources pour encourager les élèves dans cette voie, pourrait certainement obtenir de meilleurs résultats. Cela pourrait nourrir un futur travail réflexif sur l'importance de l'auto-évaluation et sur l'effort cognitif nécessaire pour favoriser l'apprentissage. Cela permettrait non seulement de sensibiliser les élèves à l'importance de cette pratique, mais aussi de les encourager à devenir des apprenants plus autonomes.

Le début de l'expérience a été très prometteur, avec des résultats encourageants. Malgré des notes globalement plus basses, la moyenne des notes obtenues par le groupe test lors de l'épreuve commune a dépassé celle du groupe témoin, ce qui est conforme aux résultats obtenus par Roediger et al. en 2011. Cependant, cet avantage s'est dissipé lors de l'évaluation suivante, et l'écart d'origine entre les deux groupes s'est reformé. Plusieurs facteurs peuvent expliquer ce phénomène. Premièrement, le délai imparti pour l'expérience était très court, ce qui signifie que certaines notions abordées dans le deuxième devoir n'ont pas pu être répétées plusieurs fois, voire pas du tout pour une partie du devoir. De plus, les principaux effets de ce dispositif peuvent se manifester sur le long terme, ce qui souligne l'importance de poursuivre ce travail dans l'espoir d'observer des bénéfices durables. Deuxièmement, les élèves du groupe témoin ont eu accès aux tests réalisés par les élèves du groupe test. J'ai fait le choix de leur donner cet accès pour des raisons éthiques. De plus, après discussion avec des élèves, certains affirment les avoir réalisés de leur côté. Malheureusement, je n'ai pas relevé le nombre d'élèves ayant effectué ces tests. Troisièmement, le choix des spécialités peut également avoir joué un rôle. Le groupe ayant un plus grand nombre de spécialités scientifiques a probablement continué à progresser dans les raisonnements scientifiques.

Les résultats concernant le stress sont très intéressants. Premièrement, on observe que le niveau de stress lors des tests au début de chaque séance ne fait qu'augmenter, mais reste néanmoins raisonnable. En revanche, le stress lié aux évaluations sommatives diminue pour le groupe test. De plus, l'écart-type sur le niveau de stress est également en légère baisse. Ces deux facteurs indiquent que le niveau de stress de l'ensemble des élèves du groupe test converge vers un niveau de stress modéré. Ces résultats semblent particulièrement intéressants à mettre en relation avec la loi de Yerkes-Dodson (1908), qui indique l'importance d'un niveau minimal de stress pour atteindre le seuil de performance optimal. Les données sur le stress du groupe témoin sont tout aussi intéressantes, l'observation d'un pic de stress lors de l'épreuve commune est associé à une chute des performances, cela est en accord avec la loi de Yerkes-Dodson. En effet, lorsque le niveau de stress est trop élevé, la performance cognitive est généralement amoindrie.

Il est important de trouver un équilibre optimal en termes de niveau de stress pour favoriser les performances académiques. Il est à mon avis trop tôt pour tirer des conclusions définitives sur l'impact du dispositif sur le niveau de stress des élèves. Cependant, les premiers résultats montrant une baisse du niveau de stress dans le groupe test sont très encourageants. Cela suggère que le dispositif pourrait avoir un effet positif sur la gestion du stress des élèves, ce qui est une information prometteuse. Ces premiers résultats obtenus sont en accord avec ceux obtenus par Brown et al. en 2016

Toutes ces données sur le stress restent néanmoins subjectives, car il était demandé à chaque élève individuellement d'évaluer son niveau de stress. Ce qui pourrait être perçu comme "un peu de stress" pour un élève peut avoir une tout autre signification pour un autre, qui pourrait le qualifier de "beaucoup de stress". De plus, certains élèves peuvent être habitués au stress à travers des activités sportives ou compétitives et avoir appris à le gérer, tandis que d'autres peuvent être en difficulté face à la gestion de ce stress. Le stress est une expérience subjective et les réponses des élèves peuvent varier en fonction de leurs antécédents, de leur personnalité et de leur environnement. Dans ce contexte, il est pourrait être intéressant de recueillir des données qualitatives supplémentaires, telles que des entretiens individuels ou des questionnaires approfondis sur les antécédents, les activités extérieures, et d'autres facteurs pouvant influencer la gestion du stress de chaque élève et leurs stratégies de gestion du stress.

Malheureusement, les résultats de la méthode actuelle n'ont pas été concluants pour réintégrer les élèves qui semblent désintéressés par l'enseignement scientifique. Cependant, il est possible d'explorer d'autres approches pour les évaluations. Envisager des évaluations où le

succès serait atteint par l'ensemble de la classe pourrait être bénéfique. Ce type d'évaluation mettrait l'accent sur la réussite ce qui pourrait aider à renforcer le sentiment d'efficacité personnelle des élèves. Les élèves pourraient retrouver confiance en leurs capacités et être encouragés à poursuivre leurs efforts dans l'apprentissage des sciences. Cette approche pourrait également aider à réduire le stress lié aux évaluations.

Biais expérimentaux :

Je n'ai pas eu la possibilité de sélectionner mon échantillon afin que mon groupe témoin ait le même niveau de départ que mon groupe test cela engendre un biais non négligeable lors de l'analyse des résultats.

Étant donné que j'ai mis en place le dispositif expérimental dans mes propres classes, je fais partie intégrante de la situation étudiée, cela peut introduire quelques biais dans cette étude. Je fais référence au risque de partialité. En d'autres termes j'aurais pu être tenté de rechercher des résultats qui confirment mes croyances ou mes pratiques professionnelles, ou de minimiser les résultats qui les contredisent. Je suis conscient de la possibilité de ce biais dans cette étude en raison de mon rôle dans cette recherche et je me suis donc forcé à rester le plus objectif possible tout au long de l'expérience.

Il y a également un biais lié à la petite taille de l'échantillon que j'ai pu tester, également appelé erreur d'échantillonnage. Elle peut survenir lorsqu'une étude ne recueille des données que sur un petit groupe de personnes. Lorsqu'un petit échantillon est utilisé pour tirer des conclusions sur une population plus large, il y a un risque que les résultats ne soient pas représentatifs de la population dans son ensemble. Les échantillons plus petits ont tendance à avoir plus de variabilité dans les données recueillies, ce qui peut entraîner des résultats moins fiables et moins précis. De plus, il peut y avoir des caractéristiques uniques dans un petit échantillon qui ne se retrouvent pas dans la population générale d'un établissement scolaire, ce qui peut conduire à des conclusions erronées. Comme cité plus haut, je n'ai pas eu la possibilité de sélectionner mon échantillon. De ce fait, il n'a donc pas forcément été le plus représentatif de la population dans son ensemble.

Les élèves absents lors de certains tests ou évaluations n'ont pas pu les rattraper, ce qui peut introduire un biais, d'autant plus, compte tenu de la petite taille de l'échantillon.

La thématique des tests a été très restreinte, il semblerait donc optimiste de généraliser ces résultats à l'ensemble de la matière.

La correction manuelle des tests de récupération peut introduire des biais potentiels. Je suis conscient de la possibilité de ce biais dans cette étude mais comme à chaque correction d'évaluation, faisant partie de cette étude ou non je prends en compte ces facteurs afin de garantir une évaluation la plus juste et équitable des élèves.

En tant qu'enseignant débutant, ma pratique professionnelle évolue à chaque séance, ce qui peut avoir un impact sur la mise en œuvre et les résultats de l'étude.

Réponse à la problématique :

La problématique de départ porte sur l'efficacité des évaluations brèves et à faible enjeu en début de séance pour améliorer les apprentissages des élèves et réduire leur stress lié aux évaluations. En raison du très court délai d'expérimentation et des nombreux biais dans cette il est difficile de tirer des conclusions cependant on peut essayer d'apporter quelques éléments de réponses à cette problématique.

Il ressort que les évaluations brèves en début de séance peuvent avoir un impact positif sur le niveau de stress des élèves. Les élèves ont tendance à ressentir moins de pression et le stress est moins variable d'une évaluation sommative à l'autre lorsqu'ils sont confrontés régulièrement à des évaluations même si celle-ci n'ont qu'un faible enjeu.

L'aspect le plus contestable vis-à-vis des résultats de l'expérience est que les évaluations brèves sont bénéfiques pour l'apprentissage des élèves. Il convient de noter que les résultats de l'expérience ont également révélé certains défis et limitations associés à cette approche. Certains élèves peuvent ne pas accorder d'importance à ces évaluations, ce qui suggère que d'autres facteurs, tels que le désintérêt pour la matière, peuvent également influencer leur engagement et leurs performances. De plus, le dispositif mis en place n'a pas toujours réussi à réengager les élèves qui semblent mettre de côté l'enseignement scientifique. Des évaluations conçues de manière à renforcer le sentiment d'efficacité personnelle et à favoriser l'engagement actif des élèves dans le processus d'apprentissage, pourraient être plus bénéfiques pour le réengagement des élèves ayant abandonné les matières scientifiques. Des évaluations simples et accessibles peuvent permettre aux élèves de réussir, ce qui peut renforcer leur confiance en leurs capacités et les encourager à s'investir davantage dans l'apprentissage des sciences.

CONCLUSION

L'état actuel de cette recherche ne permet pas de tirer des conclusions définitives sur les bénéfices des évaluations brèves et à faible enjeu en début de séance. Cependant, cette étude m'a permis de prendre conscience que les effets de l'évaluation vont bien au-delà de la simple détection des lacunes dans l'apprentissage et de l'encouragement des étudiants à travailler plus régulièrement.

Les évaluations régulières m'ont fourni un retour d'information rapide, essentiel pour corriger les lacunes dans mon enseignement et ajuster mes séances en conséquence. Ce retour d'information est crucial pour mon développement professionnel en tant qu'enseignant débutant.

Malheureusement, la mise en place de ce dispositif dans une classe de physique-chimie au niveau secondaire n'a pas permis de retrouver des résultats aussi positifs que ceux observés par Roediger et al. en 2011. Cette étude est trop courte pour observer pleinement ces effets, notamment les effets sur la mémoire à long terme.

Pour maximiser les avantages de cette approche, il est essentiel de prendre en compte les besoins et les caractéristiques individuels des élèves, ainsi que de fournir un soutien supplémentaire pour surmonter les obstacles éventuels à l'engagement et à la réussite dans l'enseignement scientifique. C'est sur ce point que j'ai certainement le moins réussi. Par conséquent, je compte poursuivre mon expérimentation mais dorénavant sans tenir compte de la note et en proposant des sujets d'évaluation différenciés afin que chaque élève puisse avoir une évaluation à son niveau. Ni trop difficile, ce qui pourrait l'empêcher d'en tirer des bénéfices, ni trop facile, car la mémorisation nécessite un effort cognitif.

En définitive, bien que les évaluations brèves et à faible enjeu en début de séance puissent contribuer à réduire le stress lié aux évaluations et à améliorer les apprentissages des élèves dans une certaine mesure, leur efficacité dépend de divers facteurs, plus particulièrement la manière dont elles sont conçues et intégrées dans le processus d'enseignement et d'apprentissage.

BIBLIOGRAPHIE

- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning : Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 940-945. <https://doi.org/10.1037/a0029199>
- Bandura, A. (2003). A. Bandura. Auto-efficacité. Le sentiment d'efficacité personnelle. *L'Orientation scolaire et professionnelle*, 33/3, 475-476.
<https://doi.org/10.4000/osp.741>
- Brown Peter C, Roediger Henry L, McDaniel Mark A, Randon-Furling Julien, & Viguiers Williams Anne. (2016). *Mets-toi ça dans la tête ! : Les stratégies d'apprentissage à la lumière des sciences cognitives / Peter C. Brown, Henry L. Roediger, Mark A. McDaniel ; préface d'Elena Pasquinelli ; traduit de l'anglais par Julien Randon-Furling et Anne Viguiers*. Éditions Markus Haller.
- Callender, A. A., & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, 34(1), 30-41.
<https://doi.org/10.1016/j.cedpsych.2008.07.001>
- Cassady, J. C. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and Instruction*, 14(6), 569-592.
<https://doi.org/10.1016/j.learninstruc.2004.09.002>
- Jeffrey, D. (2015). *Rites scolaires et identité d'élève*.
- Karpicke, J. D., & Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 704-719.
<https://doi.org/10.1037/0278-7393.33.4.704>
- 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>

Roediger, 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>

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>

Stordeur Joseph. (2014). *Comprendre, apprendre, mémoriser : Les neurosciences au service de la pédagogie / Joseph Stordeur*. De Boeck.

Viau, R. (1995). L'état des recherches sur l'anxiété en contexte scolaire. *Nouveaux cahiers de la recherche en éducation*, 2(2), 375. <https://doi.org/10.7202/1018209ar>

ANNEXES

Nom :

Prénom.....

• **Je suis stressé(e) par cette évaluation :**

Pas du tout



Presque pas



Un peu



Oui



Oui beaucoup



Calculer la température de surface de l'étoile Aldébaran.

Le profil de cette étoile est reproduit ci-contre :

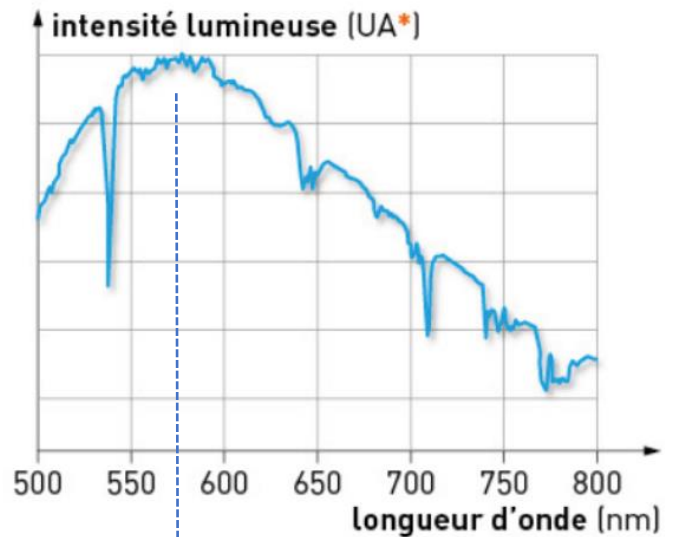
Rappel de la loi de Wien : $\lambda_{max} \times T = 2,898 \times 10^{-3}$

Avec : λ_{max} en mètre et T en Kelvin

$$\lambda_{max} = 575 \text{ nm} = 575 \times 10^{-9} \text{ m}$$

$$\lambda_{max} \times T = 2,898 \times 10^{-3}$$

$$T = \frac{(2,898 \times 10^{-3})}{\lambda_{max}} = \frac{(2,898 \times 10^{-3})}{575 \times 10^{-9}} = 5040 \text{ K}$$



575

Nom :

Prénom.....

• Je suis stressé(e) par cette évaluation :

Pas du tout



Presque pas



Un peu



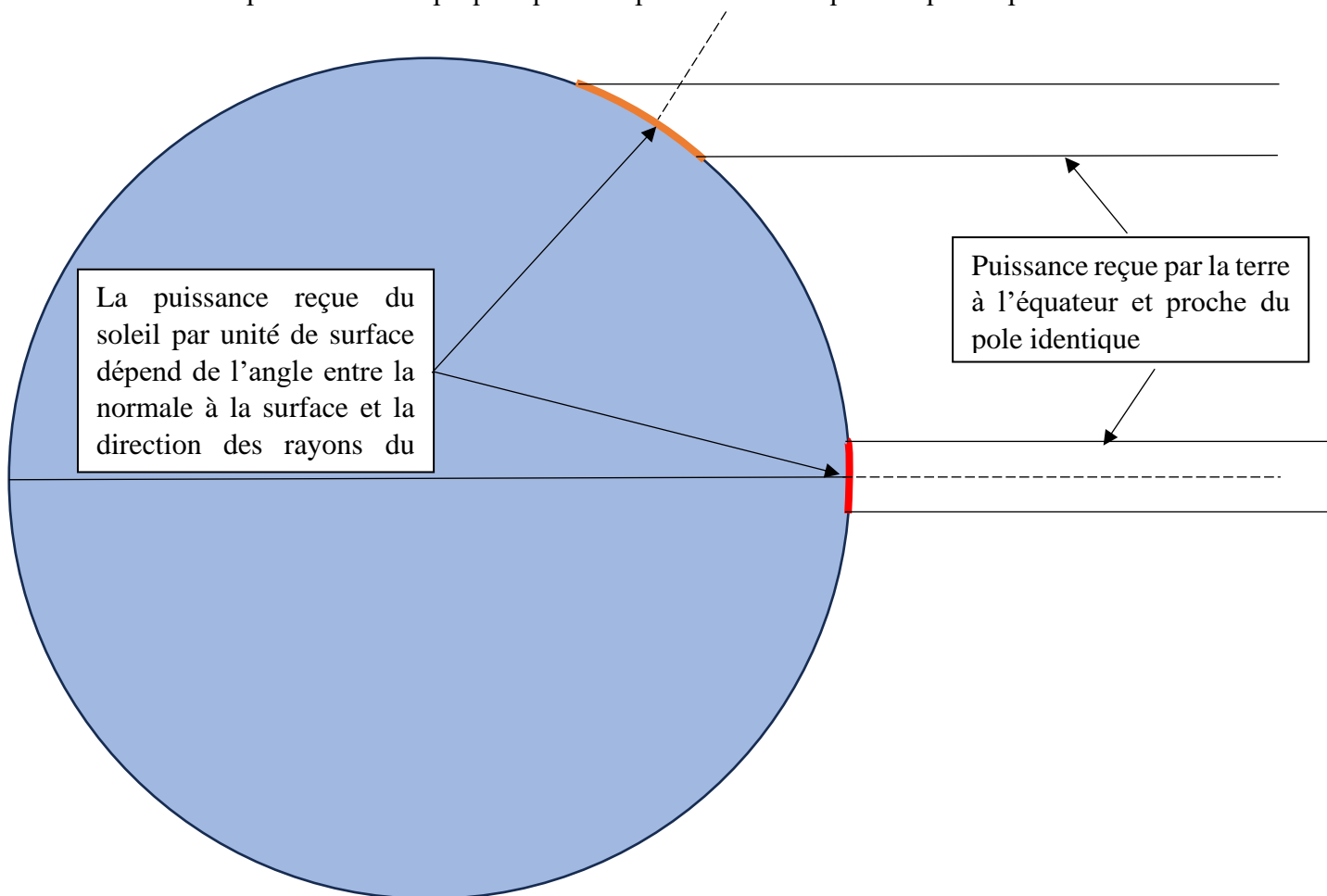
Oui



Oui beaucoup



Réaliser un schéma permettant d'expliquer qu'il fait plus chaud à l'équateur qu'aux pôles.



• Je suis stressé(e) par cette évaluation :

Pas du tout



Presque pas



Un peu



Oui



Oui beaucoup



Question 1 :

Il fait plus chaud à l'équateur qu'au pôle car :



- A. La Terre est plus proche du Soleil à l'équateur qu'au pôles.
- B. L'angle entre la normale à la surface et la direction du soleil est plus important au pôle.
- C. La banquise refroidit l'air ambiant.
- D. La réponse D.

Question 2 :

La puissance solaire reçue par unité de surface terrestre dépend :



- A. De l'heure.
- B. Du moment de l'année.
- C. De la latitude.
- D. De la distance avec le Soleil.

Question 3 :

Un méridien est :



- A. Angle mesuré à partir du méridien de Greenwich.
- B. Angle mesuré à partir de l'équateur.
- C. Un cercle qui passe par les deux pôles.
- D. L'intersection de la sphère terrestre et d'un plan parallèle à celui de l'équateur.

Question 4 :

La longueur d'un méridien terrestre est :



- A. 40 000 km
- B. 6500 km
- C. 790 km
- D. 5 000 km

Question 5 :

L'arc du méridien entre Dunkerque et Barcelone est de $9,674^\circ$. Il mesure :

- A. 40 000 km
- B. 6500 km
- C. 870 km
- D. 1075km

Question 1	B
Question 2	A B C
Question 3	C
Question 4	A
Question 5	D

- $9,674^\circ \rightarrow$ Distance Dunkerque et Barcelone
- $360^\circ \rightarrow L_m = 40\,000\text{km}$

Nom :

Prénom.....

• Je suis stressé(e) par cette évaluation :

Pas du tout



Presque pas



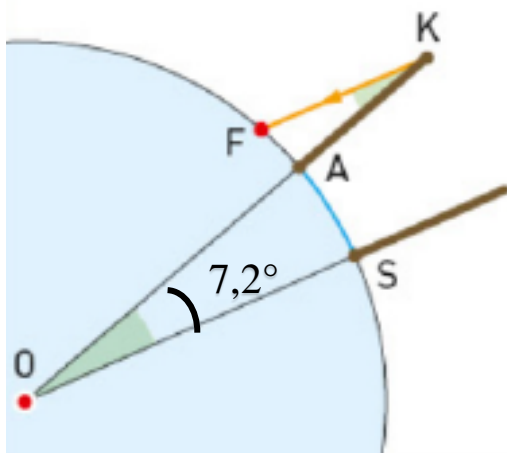
Un peu



Oui



Oui beaucoup



1. Calculer la longueur du méridien terrestre en km.

Données :

- Longueur entre Syène et Alexandrie : 800km
- Angle de l'arc de méridien entre Syène et Alexandrie : 7,2°

- 7,2° → 800 km (Distance Alexandrie Siène)
- 360° → L_m

$$L_m = \frac{800 \times 360}{7,2} = 40\,000\text{km}$$

Nom :

Prénom.....

• Je suis stressé(e) par cette évaluation :

Pas du tout



Presque pas



Un peu



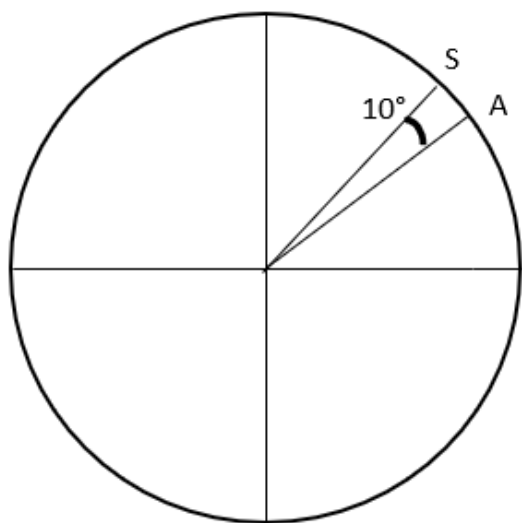
Oui



Oui beaucoup



Calculer la longueur de l'arc de méridien reliant Alger (point A) et Semur en Auxois (point B) de coordonnées géographiques 04° Est – 37° Nord et 04°Est – 47° Nord.



$10^\circ \rightarrow AS$ (Distance Alger Semur en Auxois)

$360^\circ \rightarrow L_m$

$$AS = \frac{10 \times 40000}{360} = 1\ 100\ km$$

Questionnaire 1 :

NOM Prénom :

Classe :

Je relis mes cours une ou plusieurs fois pour les apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je me fais mes propres petites évaluations pour m'aider à apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je me fais mes propres petites évaluations pour m'aider à apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Comprendre un phénomène de physique chimie signifie que je l'ai appris :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je suis stressé(e) à l'approche d'une évaluation non notée :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je suis stressé(e) à l'approche d'une évaluation notée :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Questionnaire 2 :

NOM Prénom :

Classe :

Je relis mes cours une ou plusieurs fois pour les apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je me fais mes propres petites évaluations pour m'aider à apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je me fais mes propres petites évaluations pour m'aider à apprendre :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Comprendre un phénomène de physique chimie signifie que je l'ai appris :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je suis stressé(e) à l'approche d'une évaluation non notée :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je suis stressé(e) à l'approche d'une évaluation notée :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Les courtes évaluations à chaque séance m'ont aidé à apprendre mon cours de physique chimie :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord

Je souhaite continuer les évaluations courtes au début de chaque séance :

Pas du tout
d'accord

Plutôt pas
d'accord

Indifférent

Plutôt
d'accord

Tout à fait
d'accord