

Dans quelle mesure les objectifs d'une évaluation sommative à enjeu en mathématique peuvent-ils influencer les performances des élèves en retard selon qu'elle soit axée sur la performance ou la maîtrise ?

Auteur : Flore, Nicolas

Promoteur(s) : Fagnant, Annick

Faculté : Faculté de Psychologie, Logopédie et Sciences de l'Éducation

Diplôme : Master en sciences de l'éducation, à finalité spécialisée en formation des adultes

Année académique : 2024-2025

URI/URL : <http://hdl.handle.net/2268.2/24593>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

« Dans quelle mesure les objectifs d'une évaluation sommative avec enjeu en mathématique peuvent-ils influencer les performances des élèves en retard selon qu'elle soit axée sur la performance ou sur la maîtrise ? »

Promotrice: FAGNANT Annick

Lecteurs: MONSEUR Christian

PRESSIA Fabian

Mémoire présenté par FLORE Nicolas en vue de l'obtention du grade
de Master en sciences de l'éducation à finalité spécialisée en
formation des adultes

Année académique 2024 - 2025

Remerciements

Avant toute chose, je souhaite remercier sincèrement toutes celles et ceux qui, à leur manière, ont apporté leur soutien à l'élaboration de ce travail de fin d'études.

Dans un premier temps, je tiens à remercier ma promotrice, Madame Annick Fagnant, Professeure à la faculté de Psychologie, Logopédie et Sciences de l'Éducation de l'ULiège, pour l'accompagnement et l'encadrement dispensés durant la mise en place puis la réalisation de ce projet. Je tiens également à remercier Monsieur Jean Baron pour les conseils et le suivi de ce mémoire.

Je remercie grandement Monsieur Christian Monseur et Monsieur Fabian Pressia pour l'attention qu'ils ont accordée à ce mémoire en acceptant d'en assurer la lecture.

En outre, je tiens à exprimer toute ma gratitude envers les directions d'école, les enseignants et les élèves qui, en acceptant de participer à ce projet, nous ont accordé une grande confiance. Mes remerciements vont aussi à toutes celles et ceux qui ont donné de leur temps et de leur énergie pour mettre à bien ce projet en se renseignant, en contactant des connaissances, des collègues, etc.

Enfin, je souhaite exprimer ma profonde reconnaissance envers mes parents, mes frères, ma famille, mes amis et mes proches qui n'ont cessé de m'aider, de m'encourager et me soutenir dans la réalisation de ce travail, mais aussi plus globalement dans l'aboutissement de ce master en Sciences de l'Éducation.

Tables des matières

1. Introduction	5
2. Revue de la littérature	7
2.1. L'évaluation vue de manière holistique	7
2.1.1. Les fonctions de l'évaluation dans notre système éducatif	7
2.1.2. Focus sur deux types d'évaluation	8
2.2. Les processus et phénomènes en jeux en situation évaluative	10
2.2.1. La motivation et les buts d'accomplissements	11
2.2.2. L'influence de la représentation de soi	14
2.2.3. La comparaison sociale	15
2.3. La menace des stéréotypes liée au redoublement	16
2.3.1. Les stéréotypes sous l'angle de la psychologie sociale	16
2.3.2. La menace du stéréotype en situation évaluative	16
2.3.3. Les redoublants face aux stéréotypes	17
2.4. Le redoublement en Fédération Wallonie-Bruxelles	18
2.5. Les mathématiques	19
2.5.1. Une discipline pas anodine	19
2.5.2. La question du genre en mathématique	20
2.6. L'anxiété liée aux mathématiques	21
2.6.1. Définition et situation en Fédération Wallonie-Bruxelles	21
2.6.2. Les sources de l'anxiété	22
2.6.3. Les conséquences de l'anxiété	23
2.7. Le concept de soi	25
2.7.1. Définition et caractéristiques du concept de soi	25
2.7.2. Distinguer le concept de soi et d'autres concepts liés	26
2.7.3. Concept de soi et performance	27
2.7.4. Focus sur deux catégories d'élèves	28
2.8. Conclusion de la revue de la littérature	29

3. Question de recherche et hypothèses	30
4. Méthodologie	35
4.1. Méthode choisie	35
4.2. Description de l'échantillon	36
4.3. Procédure générale	39
4.4. Instruments de mesure	42
4.4.1. Test de mathématiques	42
4.4.2. Questionnaire contextuel	43
5. Présentation, analyse et discussion des résultats	44
5.1. Consistance interne des instruments de mesure	44
5.1.1. Validation du questionnaire contextuel	44
5.1.2. Validation du test de mathématique	46
5.2. Résultats du test de mathématique	48
5.2.1. Une analyse préalable de nos échantillons	48
5.2.2. Comment qualifier la différence de moyenne entre les échantillons ?	50
5.2.3. Hypothèses 1a, 1b, 1c et 1d	51
5.3. Résultats du questionnaire contextuel	53
5.3.1. Hypothèses 2a, 2b et 2c	55
5.3.2. Hypothèses 3a, 3b et 3c	55
5.3.3. Hypothèses 4a, 4b et 4c	56
5.4. Interprétation et discussion des résultats	57
6. Limites et perspectives	60
7. Conclusion	61
Bibliographie	63
Table des tableaux	67
Table des annexes	68

1. Introduction

« *Comment a-t-il pu encore rater son évaluation ? Comment mon élève, mon enfant, ma sœur, mon ami... a-t-il bien pu échouer à son contrôle ?* » Voici quelques questions que vous devriez déjà vous être posées une, voire de nombreuses fois concernant un(e) proche. Ces questions sont légitimes, car elles cherchent à déterminer les causes d'un échec. Elles interrogent une dimension parfois sous-estimée à l'école, mais bien présente dans notre enseignement : l'évaluation. Les réponses les plus naturelles à ces questions semblent prendre racine dans notre société où il existe une croyance commune très ancrée qui consiste à justifier les échecs des élèves aux évaluations uniquement par leurs compétences et leurs capacités intrinsèques (Toczek & Souchal, 2017). Toczek et Souchal (2017), ainsi que d'autres auteurs avant elles, s'efforcent de déconstruire cette vision innéiste et démontrent qu'une série de facteurs issus du contexte évaluatif peuvent être une source de menace pour certains élèves et donc également jouer un rôle dans les performances scolaires. En effet, depuis un certain nombre d'années, diverses études montrent à quel point la situation évaluative dans laquelle les élèves performant est cruciale (Lafontaine & Toczek, 2023). En réalité, toute une série de phénomènes psychosociaux peut causer des différences de performances (Lafontaine & Toczek, 2023). Monteil et Huguet (2011) indiquent par exemple que le type d'évaluation, la motivation de l'élève, les interactions sociales, les expériences passées, etc. jouent un rôle dans l'accomplissement d'une tâche évaluative et qu'il faut y prêter attention.

De plus, un autre phénomène bien connu et largement étudié par la psychologie sociale semble également apparaître en situation évaluative : la menace du stéréotype. L'étude de ce mécanisme montre que des signaux du contexte évaluatif peuvent dans certains cas faire réussir les élèves et dans d'autres les faire échouer selon que ceux-ci aient subi la menace du stéréotype ou non (Lafontaine & Toczek, 2023). Il nous semble crucial de prendre conscience des contextes évaluatifs qui augmentent ou diminuent cette menace du stéréotype et qui ont des répercussions conséquentes sur les performances des élèves.

Ces réflexions nous ont donc amenés à nous poser la question suivante : « dans quelle mesure l'objectif d'une évaluation sommative avec enjeu peut-il influencer les performances des élèves en retard selon qu'elle soit axée sur la performance (c'est-à-dire une évaluation ayant pour but de principalement comparer et sélectionner les élèves) ou qu'elle soit axée sur la maîtrise (c'est-à-dire une évaluation présentée comme ayant également un but d'apprentissage) ? ».

L'objectif central de cette recherche est d'observer si l'effet de la « menace du stéréotype doublant » suggéré par le contexte évaluatif se marque davantage dans un contexte d'évaluation sommative axée sur la performance comparativement à un contexte d'évaluation sommative axée sur la maîtrise. Concrètement, nous souhaitons manipuler le contexte évaluatif de deux façons comme le font Toczek et Souchal (2017) dans leur étude afin de voir si certaines manières de présenter les évaluations sommatives peuvent être davantage bénéfiques pour les élèves en retard scolaire et pour d'autres catégories d'élèves. Ensuite, différents sous-objectifs découlent du principal cité précédemment. En effet, dans un second temps, le but est de mieux comprendre cet effet de « menace du stéréotype du doublant » selon différentes caractéristiques comme le genre, les buts poursuivis, le concept de soi en mathématique ou l'anxiété liée à cette matière. Nous tenterons par exemple de voir si les doublants poursuivent davantage des buts de performance-évitement ou non, si les filles sont affectées davantage par le contexte évaluatif selon qu'elles soient doublantes ou non, etc. De façon plus secondaire, ce qui motive également ce travail, c'est entre autres la volonté de faire réfléchir les lecteurs de cette recherche et particulièrement les enseignants à la portée et aux conséquences de la façon de présenter les enjeux d'une évaluation sur les élèves en retard scolaire.

Afin de tester nos diverses hypothèses, nous avons pu mettre en place notre dispositif dans vingt-quatre classes de troisième secondaire de l'enseignement général et technique de transition de la Fédération Wallonie-Bruxelles (FW-B). Notre dispositif comprend une évaluation sommative à enjeu de mathématiques sur les fonctions. Celle-ci a été mise en place dans deux conditions expérimentales que nous comparerons. En effet, cette évaluation a d'abord été conduite dans douze classes en mettant en avant des enjeux de performance et en parallèle, dans douze autres classes en insistant cette fois-ci sur des enjeux de maîtrise. Ces deux manières de présenter cette même évaluation correspondent aux deux conditions de notre étude.

En outre de ce test de mathématique, un questionnaire contextuel sur le concept de soi, l'anxiété en mathématiques et les buts d'accomplissements a également été soumis à tous les élèves. Nous détaillerons davantage le contenu de l'entièreté du dispositif dans la partie méthodologie. Nous développerons également dans cette partie la manière dont nous avons constitué et appareillé nos échantillons.

2. Revue de la littérature

2.1. L'évaluation vue de manière holistique

Avant d'observer plus en détail les phénomènes qui se déroulent en situation évaluative, il nous semble important de prendre un peu de hauteur et de regarder les choses de manière holistique. En effet, pour bien comprendre ce qu'il se passe au niveau des individus, des élèves, il est utile d'avoir conscience des enjeux macro qui se jouent en arrière-plan.

2.1.1. *Les fonctions de l'évaluation dans notre système éducatif*

Tout d'abord, il est important de prendre conscience de la place qu'occupe l'évaluation dans notre système éducatif, car l'un et l'autre sont liés. Il semble que notre système éducatif tente de remplir deux fonctions/missions. En effet, dans un premier temps, l'école se veut accessible à tout un chacun, quelle que soit l'origine géographique ou socio-économique, et cela dans le but de transmettre des savoirs et de faire grandir les jeunes. Butera (2023) parle de fonction de formation tandis qu'Autin et ses collaborateurs (2019) évoquent une fonction éducative, mais l'idée reste la même, l'école forme et éduque tous les enfants. Parallèlement, l'école remplit également une fonction de sélection. Florin et ses collaborateurs (2023) indiquent que dès le dix-neuvième siècle, l'école sélectionne « pour former les futures élites » (p. 31). De nos jours encore, pour remplir sa fonction de sélection, l'école va recourir à des évaluations qui vont avoir « un impact sur le cursus scolaire, sur les opportunités de choix de filières d'études ou de formation, et in fine sur l'emploi et les positions sociales auxquels on peut accéder en fonction des diplômes obtenus » (Darnon et al., 2009, cités par Butera, 2023, p. 5). Il est intéressant de noter que ces fonctions de l'école se traduisent dans le choix et la conception des différentes évaluations organisées (Florin et al., 2023). Autrement dit, certains types d'évaluations tendent à remplir plus ou moins fortement l'une ou l'autre fonction (de formation ou de sélection). Selon Dornbush et ses collaborateurs (1996), la société a conscience de ces fonctions (Dornbush et al., 1996, cités par Butera, 2023). Cependant, il apparaît que ces fonctions sont contradictoires, ce qui force l'école à gérer un paradoxe (Florin et al., 2023).

Face à ce dilemme, le système éducatif et les enseignants doivent faire des choix en termes d'évaluations. En réalité, il semble que ce soient les évaluations au service de la sélection qui soient très majoritairement pratiquées (Florin et al. 2023). En effet, l'étude « Teaching and Learning International Survey » (TALIS) de 2018 a cherché à quantifier les types d'évaluations utilisées dans les pays de l'OCDE. Il apparaît que l'évaluation sommative/normative soit la plus

majoritairement pratiquée (Pasquini, 2023). Ce type d'évaluation permet en effet de classer, comparer ou hiérarchiser les élèves entre eux au travers de notes par exemple.

2.1.2. Focus sur deux types d'évaluation

Il est donc possible d'entreprendre des évaluations au service des apprentissages pour tous ou des évaluations qui permettent de sélectionner les plus performants. Comme expliqué ci-dessus, les évaluations pratiquées en classe vont favoriser la poursuite des fonctions de sélection ou de formation. La littérature distingue et différencie de nombreux types et d'appellations différentes concernant les évaluations. Comme esquissé précédemment, dans le cadre de notre étude, nous mettons en place une évaluation dite sommative à enjeu. Cette évaluation sommative intervient dans deux conditions différentes et est donc présentée de deux manières différentes ; l'une dirigée vers un axe normatif et l'autre vers un axe formatif. Par conséquent, ces deux types d'évaluation nous intéressent particulièrement, c'est pourquoi nous nous concentrerons essentiellement sur celles-ci. Distinguer ces deux types d'évaluations (dites « formatives » et « normatives ») nous permettra de comprendre davantage les enjeux et les phénomènes qui y sont reliés.

L'évaluation présentée comme celle qui permet de poursuivre le but de formation est l'évaluation dite formative. Dierendonck et Cavaco (2017) la qualifient d'évaluation « pour apprendre », autrement dit, qui est au service de l'apprentissage. Marc Romainville (2017) explique que l'évaluation formative permet pas de prendre une décision, si ce n'est des décisions régulatrices. En effet, elle offre la possibilité aux élèves et à l'enseignant de prendre connaissance de ce qui est acquis et ce qui ne l'est pas au moment du test. Cette prise d'information permet, durant l'apprentissage, « de réaliser en continu des ajustements des activités d'enseignement (pour les enseignants) et d'apprentissage (pour les élèves) » (Florin et al., 2023, p. 31). Florin et ses collaborateurs (2023) rajoutent, sur base de diverses méta-analyses, que l'évaluation formative apporte plusieurs bénéfices dont une amélioration des apprentissages des élèves et notamment de ceux dits en difficulté ou moins performants. Bien que les évaluations formatives et sommatives aient parfois été opposées, certains chercheurs s'emploient depuis quelques années à les rapprocher. Lavault et Allal (2016) ont par exemple développé le principe d'Évaluation soutien d'apprentissage (ESA) (Lavault & Allal, 2016, cités par Fagnant, 2023). Ce concept invite à dépasser la vision première et parfois stéréotypée de l'évaluation formative. Ces auteurs offrent une nouvelle perspective sur les ponts qui peuvent être bâtis entre l'évaluation formative et l'évaluation sommative. Leur modèle permet d'envisager également l'évaluation sommative comme pouvant aussi être une forme

d'évaluation pour apprendre, une évaluation qui soutient l'apprentissage. De cette manière, ces chercheurs démontrent que l'évaluation sommative et l'évaluation formative ne sont pas incompatibles. Cela va dans le sens de notre présente recherche puisque nous émettons l'hypothèse qu'une évaluation sommative combinée à des éléments formatifs permet à certaines catégories d'élèves d'être davantage performants.

De son côté, l'évaluation normative semble être celle qui entretient le plus la fonction de sélection de l'école. Deutsch (1979) affirme que « ce lien [entre fonction de sélection et évaluation normative] serait basé sur la capacité de l'évaluation normative à attribuer une valeur comparative au travail de l'élève et donc du mérite à l'élève même, mérite qui ensuite est rendu tangible par des récompenses académiques telles que la position hiérarchique dans la classe, la promotion, ou l'orientation vers des filières désirables » (Deutsch, 1979, cité par Butera, 2023, p. 6).

Ce lien étant avéré, il est utile d'analyser plus en profondeur la question de l'évaluation normative. Butera (2023) donne sa définition de l'évaluation normative qui consiste selon lui à « attribuer une valeur à la production des élèves, qu'elle soit représentée par un nombre, une lettre, un jugement, ce qui rend le rang de cette production (et de l'élève) visible dans la comparaison à un standard ou au reste de la classe » (p. 6). La nature de l'évaluation normative est dans la plupart des cas sommative. C'est-à-dire que sur base de la réussite ou non d'un certain nombre d'items, une note globale est attribuée (ex : contrôles, examens, etc.). Dierendonck et Cavaco (2017) la qualifient d'évaluation « pour contrôler », autrement dit, qui permet de certifier ou non, au terme d'un apprentissage, les acquis des élèves. Une des caractéristiques principales de l'évaluation normative est qu'elle compare la performance de l'élève à celle des autres élèves et non pas à des objectifs définis préalablement comme c'est le cas de l'évaluation critérielle (Glaser, 1963 ; de Landsheere, 1992, cités par Crahay, 2019). En effet, la simplicité et la clarté d'une note attribuée à un élève permettent très aisément, notamment aux enseignants, mais aussi aux élèves de se comparer (Butera, 2023). De cette comparaison peut très naturellement découler un classement et une hiérarchisation au sein de la classe.

Il semble donc que les enseignants aussi doivent faire face à un dilemme de taille. Ceux-ci doivent choisir entre proposer des évaluations qui permettent aux élèves de maîtriser la matière et proposer des évaluations qui permettent de récompenser les plus performants en leur attribuant de « bonnes » notes, voire des diplômes (Souchal et al., 2014). Néanmoins, de nouvelles perspectives ont été dégagées par certains auteurs afin de dépasser la vision

dichotomique qui oppose formatif et sommatif/normatif. C'est le cas du concept d'Evaluation comme soutien d'apprentissage qui relie ces deux formes d'évaluation (Lavault & Allal, 2016, cités par Fagnant, 2023).

Cependant, il apparaît que les enseignants ont recours de façon très majoritaire à l'évaluation normative, et ce pour des raisons historiques, politiques et économiques (Florin et al., 2023 ; Dierendonck & Cavaco, 2017). Ce modèle méritocratique qui domine dans notre enseignement permet à l'évaluation normative de récompenser les étudiants considérés comme les plus méritants par de meilleures notes. Selon Florin et ses collaborateurs (2023), ce système de note « traduit de façon lisible une hiérarchie entre les élèves » (p. 31). Selon Rey et Feyfant (2014), sortir de cette tradition fortement ancrée de l'évaluation à référence normative demande un investissement conséquent et majeur de la part des enseignants. Ils rajoutent qu'évaluer pour apprendre équivaut à « se situer à rebours d'une tradition où l'évaluation est conçue dans un modèle cybernétique comme un contrôle a posteriori, quasiment technique, d'un apprentissage qui serait indépendant des modalités d'évaluation » (Rey & Feyfant, 2014, cités par Dierendonck & Cavaco, 2017, p 100). Nous observons donc qu'une série de situations évaluatives peuvent être mises en place. Or, comme nous le verrons par la suite, ces diverses situations évaluatives entraînent des contextes, des phénomènes et des effets fort différents.

2.2. Les processus et phénomènes en jeux en situation évaluative

Une question qui nous semble pertinente de se poser à cette étape est de savoir si tous les élèves sont égaux face à ces situations évaluatives. En effet, l'objectif de notre étude étant d'observer si les élèves en retard performant de façon différentielle ou non en situation d'évaluation sommative axée sur la performance (davantage normative) et axée sur la maîtrise (davantage formative, dite ESA), il est utile de s'intéresser aux recherches déjà effectuées sur le sujet. En réalité, il apparaît que les évaluations normatives amènent « à contribuer à la reproduction des inégalités scolaires déjà existantes, notamment les inégalités de classe sociale » (Butera, 2023, p. 8). À partir de ce constat, il est naturel de chercher à comprendre l'origine, mais aussi les hypothèses explicatives de ces inégalités en situation évaluative.

Toczek et Souchal (2017) se sont intéressées à la façon dont les écarts de performances et les échecs de certains groupes sociaux étaient justifiés. Selon elles, les arguments cités pour justifier ces échecs sont principalement des facteurs intrinsèques aux élèves. Elles expliquent que ce sont les capacités de l'élève qui sont remises en cause lors d'un échec. Elles dénoncent

ce raisonnement selon lequel « ce sont les caractéristiques de l'élève qui détermineraient leur plus ou moins grande réussite dans une tâche » (p.22). Ce type d'explication qualifié d'« innéiste » commence, depuis quelques années, à laisser sa place à une autre hypothèse explicative ; celle de la psychologie sociale. En effet, cette nouvelle voie tente de montrer qu'il ne faut pas négliger l'impact des situations évaluatives qui sont tout aussi déterminantes que les capacités propres des individus (Toczek & Souchal, 2017). Monteil et Huguet (2001) vont dans le même sens. Ils rajoutent que le contexte social, lors d'une évaluation, fait partie intégrante du fonctionnement cognitif des individus. Finalement, cette nouvelle voie explicative nous amène à nous détacher de la croyance commune qui consiste à penser que lors de la passation d'un test, l'enfant est seul, dans une bulle et que rien ni personne n'a d'influence sur lui. Il semble plutôt que le contexte social puisse avoir des effets sur le fonctionnement cognitif des élèves en situation d'évaluation. Nous allons donc par la suite nous attarder sur les découvertes de la recherche concernant ces effets psychosociaux qui semblent être en jeu lors des situations évaluatives. Effets qui pourraient donc justifier en partie les différences de performances de certains élèves.

2.2.1. La motivation et les buts d'accomplissements

La littérature s'est penchée sur la façon dont les différents types d'évaluation, notamment l'évaluation formative et normative, pouvaient influencer la motivation, la persévérance et la performance des élèves. Diverses études sont par exemple arrivées à la conclusion que les évaluations de type normatif avaient des effets délétères sur l'intérêt des élèves à apprendre (Butera, 2023). Concrètement, « la note tuerait le plaisir d'apprendre et favoriserait les stratégies d'apprentissage superficielles permettant de performer aux évaluations pour tous les types d'élèves » (Archambault & Chouinard, 2006 ; Deci, Koestner & Ryan, 1999, cités par Lafontaine & Toczek, 2023, p. 24).

De plus, selon que l'évaluation soit normative ou formative, la représentation que se font les élèves de la tâche changerait, ce qui influencerait leur performance (Souchal & Toczek, 2010). Cette découverte, née des recherches de Nicholls (1978, 1979, 1984) et Dweck (1975), fait référence à la théorie des buts d'accomplissements (Lafontaine & Toczek, 2023). Souchal et Toczek (2010) définissent les buts d'accomplissements comme étant « les raisons pour lesquelles les individus souhaitent réussir et ce qu'ils considèrent une réussite » (p. 15). Les différentes recherches sur le sujet ont abouti à distinguer les buts de « maîtrise » et les buts de « performances ». Les buts de maîtrise correspondent au désir d'apprendre, l'individu cherche à augmenter ses connaissances et à maîtriser la tâche (Darnon, Buchs & Butera, 2006). En

revanche, les buts de performance correspondent au désir d'être meilleur que les autres (Darnon, Buchs & Butera, 2006). Néanmoins, une précision quant à ces buts de performance a été ajoutée. En effet, il faudrait distinguer les buts performance-approche (chercher à être meilleur que les autres) des buts performance-évitement (éviter d'être moins compétent que les autres) (Elliott, 1999, cité par Lafontaine & Toczek, 2023). Il est intéressant de noter que, selon Souchal et ses collaborateurs (2014), chaque type d'évaluation a tendance à favoriser la poursuite de certains buts. Elles expliquent que « les pratiques d'évaluation qui se concentrent sur les aspects normatifs augmentent les objectifs de performances, tandis que les pratiques d'évaluation qui soulignent l'importance des progrès sont susceptibles d'augmenter les objectifs de maîtrise » (Souchal et al., 2014, p. 127).

Cependant, il semble que le type d'évaluation ne soit pas le seul facteur qui détermine les buts d'accomplissements. En effet, Darnon, Buchs et Butera (2006) indiquent que le choix entre ces différents buts peut découler de facteurs situationnels (ex : les objectifs de l'évaluation, le type de tâche...), mais aussi de dispositions personnelles (ex : la conception que l'élève a de la compétence et de l'intelligence). Nicholls (1984) avance que face à une tâche qui offre un défi modéré, sans stress physique ou psychologique et sans pressions majeures de son environnement, les gens ont davantage tendance à viser des buts de maîtrise (Nicholls, 1984, cité par Sénéchal, 1997). Par contre, face à une tâche qui vise à tester les habiletés des personnes, qui induit un climat de compétition interpersonnel ou qui amène les individus à se préoccuper de l'image qu'ils renvoient de manière trop importante, ce sont des buts de performances qui sont davantage adoptés (Sénéchal, 1997).

Néanmoins, une question nous semble encore sans réponse : le fait de privilégier des buts de performances plutôt que des buts de maîtrise est-il néfaste pour les élèves ? Souchal et Toczek (2010), sur base de travaux antérieurs, ont tenté de répondre à cette question. Elles indiquent tout d'abord que les élèves qui poursuivent des buts de maîtrise sont susceptibles d'obtenir des résultats bénéfiques en termes d'apprentissage. En effet, les buts de maîtrise sont positivement corrélés avec la motivation intrinsèque (Butera, 2023). Or, la motivation intrinsèque est une des principales variables explicatives de la réussite scolaire (Hidi & Harackiewicz, 2000, cités par Souchal & Toczek, 2010). Lorsqu'ils poursuivent des buts de performance, les élèves ne voient pas l'évaluation comme une opportunité de progresser, mais comme une tâche menaçante (Souchal & Toczek, 2010). Lorsqu'ils se sentent en difficulté, ces élèves vont avoir tendance à douter d'eux-mêmes, à se décourager, à adopter des stratégies peu efficaces et donc à être moins performants (Ames, 1984 ; Dweck, 1999, cités par Souchal &

Toczek, 2010). Néanmoins, même si dans l'ensemble la poursuite de buts de performance a davantage de conséquences néfastes que les buts de maîtrise, cela n'est pas systématiquement le cas. Il apparaît par exemple que les garçons sont davantage performants lorsqu'ils ressentent une certaine compétition et lorsqu'ils poursuivent des buts de performance comme lors d'évaluations normatives (Duda, 1988, cité par Toczek & Souchal, 2017). Cette conclusion est le fruit d'une expérimentation menée par Souchal et ses collaboratrices (2014). Nous estimons pertinent de décrire brièvement cette étude, car cette présente recherche s'en inspire et s'en rapproche. Concrètement, ces chercheuses sont parties d'un constat : les filles subissent des stéréotypes négatifs les concernant en sciences, ce qui les rend moins performantes. Souchal et ses collègues (2014) ont soumis à des élèves un test de sciences présenté de trois manières différentes : axé sur la performance, axé sur la maîtrise ou évaluation formative sans notation. Leur but était de trouver une manière d'utiliser l'évaluation qui ne nuise à aucun des deux sexes. Voici les résultats de cette expérimentation : « dans la condition d'évaluation axée sur la maîtrise, les garçons et les filles ont atteint un niveau similaire, tandis que la condition d'évaluation axée sur la performance a réduit les performances des filles et que la condition d'absence d'évaluation a réduit les performances des garçons » (Souchal et al., 2014, p.125). En conclusion, cette étude a permis de montrer que présenter une évaluation de sciences comme une manière d'améliorer ses apprentissages permet d'augmenter les performances des filles sans nuire à celles des garçons. Cette présente étude, bien que quelque peu différente, poursuit un objectif semblable qui est de concevoir et présenter les évaluations sommatives d'une manière qui pourrait augmenter les performances des élèves en retard sans pour autant nuire à celles des élèves à l'heure.

Un dernier élément concernant la poursuite des buts d'accomplissements nous intéresse dans le cadre de notre recherche. Il semble que les buts de performance-approche n'aient pas les mêmes conséquences que les buts de performances-évitement. Selon Souchal et Toczek (2010), les derniers ont de manière quasiment systématique des effets délétères sur les performances des élèves. La poursuite de tels buts contraindrait les individus à adopter des stratégies d'auto-handicap, à diminuer leur compétence et donc à obtenir de plus faibles résultats (Souchal & Toczek, 2010). Lafontaine et Toczek (2023) apportent une précision. Elles déclarent qu'« annoncer aux élèves, avant une évaluation, que celle-ci va être notée renforcerait les buts de performance-évitement [pour la majorité des élèves] et pas nécessairement les buts de performance-approche, qui sont plus favorables aux apprentissages » (p. 24). Finalement, toutes ces recherches nous permettent de prendre conscience d'une série de conséquences liées

aux situations évaluatives. En fait, certaines évaluations, certains contextes évaluatifs favoriseraient la poursuite de buts d'accomplissement bien particuliers. Or, certains de ces buts ont des effets délétères sur les performances des élèves (ex : les filles moins performantes en situation évaluative sommative normative).

2.2.2. L'influence et la représentation de soi

Il semble que l'influence des situations évaluatives ne se traduit pas seulement dans la poursuite de buts d'accomplissements. En effet, Monteil et Huguet (2001) ont démontré que la dimension autobiographique et les expériences académiques passées pouvaient également influencer de manière déterminante les performances des élèves en situation d'évaluation. Toczek et Souchal (2017) parlent d'interactions entre un contexte externe (le contexte évaluatif) et interne (connaissance de soi). Pour arriver à cette conclusion, Monteil et Huguet (2001) ont réalisé une étude qui consistait à soumettre un test à des élèves. Dans la première condition, l'exercice était présenté comme une activité de géométrie tandis que dans la deuxième condition, il était présenté comme tâche de dessin. Les résultats ont permis de montrer que lorsque la tâche est présentée différemment (le contexte externe change), cela fait émerger le contexte interne différemment chez les participants également. Lorsque la tâche est présentée comme étant un test de géométrie, les élèves ayant des expériences antérieures douloureuses dans cette discipline sont moins performants. Ces mêmes élèves sont néanmoins davantage performants au même test lorsque celui-ci est présenté comme une activité de dessin. Selon Huguet (2018), cette étude permet de démontrer à quel point « les représentations de soi influencent les productions cognitives, dont on perçoit qu'elles ne sont pas réductibles aux seules capacités des auteurs » (p. 175).

Ensuite, certains auteurs ont cherché à comprendre par quel mécanisme ce phénomène se produisait. Tout d'abord, les connaissances de soi peuvent être activées au travers de signaux provenant de la situation évaluative comme le domaine de l'évaluation (ex : géométrie). L'individu va alors se rappeler d'expériences plus ou moins positives qui vont elles-mêmes être associées à des émotions diverses (Lafontaine & Toczek, 2023). Huguet (2018) explique que pour les élèves ayant des expériences douloureuses, « le souvenir très accessible de leurs échecs passés et la perspective d'un nouveau naufrage leur prennent une telle quantité d'attention qu'ils sont dans l'incapacité de se consacrer pleinement dans la tâche » (p. 177).

Enfin, certains auteurs ont mis en exergue un autre phénomène en lien avec la connaissance de soi ; la réputation d'infériorité. Selon Toczek et Souchal (2017), face à des échecs répétés

qui provoquent un sentiment d'infériorité, les élèves vont progressivement intérioriser les stéréotypes négatifs les visant. Par exemple, les élèves redoublants ont la réputation d'être endormis ou paresseux (Crisafulli et al., 2002, cités par Crahay, 2019). Autrement dit, ces élèves vont progressivement intérioriser et accepter le statut inférieur qui leur est attribué et qui par ailleurs, peut être fictif. En situation évaluative, cette réputation d'infériorité fait douter les élèves de leur niveau scolaire et de leurs capacités (Croizet, Désert, Dutrévis & Leyens, 2003). En effet, « se percevoir comme une personne ayant peu d'habiletés pour réaliser une tâche peut nuire à l'exécution des actions nécessaires à sa réalisation et, in fine, affecter sa réussite, alors que se percevoir comme compétent pour réaliser une tâche peut soutenir la mise en œuvre des comportements nécessaires à l'exécution de celle-ci et, in fine, mener à sa réussite » (Pansu, 2023, cité par Lafontaine & Toczek, 2023, p. 26). Ces diverses recherches et découvertes nous permettent de comprendre qu'en situation évaluative, la représentation de soi, le passé académique, mais aussi le sentiment d'infériorité peut être en partie responsable de faibles performances.

2.2.3. La comparaison sociale

Darnon, Buchs et Butera (2006) nous invitent à ne pas observer les effets des contextes évaluatifs uniquement du point de vue individuel de l'élève. Ils précisent qu'il ne faut pas sous-estimer la présence d'autrui en situation d'évaluation, car cette présence peut, selon eux, être source de comparaison sociale. La comparaison sociale peut être définie comme « un processus psychologique qui consiste à mettre en relation avec soi des informations de nature sociale, comme par exemple, la performance d'un ou de plusieurs autres élèves » (Wood, 1996, cité par Goudeau, 2023, p. 94). Selon Muller et Butera (2007), ce phénomène de comparaison peut apparaître dès lors que des individus sont rassemblés pour effectuer une même tâche. Ces comparaisons sont dommageables, car elles ont des effets néfastes sur l'état émotionnel ou le fonctionnement cognitif et in fine les performances des élèves (Toczek & Souchal, 2017). Concrètement, en situation évaluative, lorsque les élèves voient ou imaginent que les autres sont plus performants qu'eux, ils remettent en question leurs propres capacités à réussir et se sentent menacés (Goudeau, 2023). Lafontaine et Toczek (2023) expliquent que « la réussite des autres peut atteindre le besoin de maintenir une bonne image de soi, ce qui génère du stress et un cortège d'émotions et de pensées négatives très coûteuses en ressources attentionnelles » (p. 30). Finalement, les évaluations qui favorisent un esprit compétitif rendent la présence d'autrui menaçante et focalisent une grande partie de l'attention des élèves en difficulté (Muller & Butera, 2007).

2.3. La menace du stéréotype liée au redoublement

2.3.1. Les stéréotypes sous l'angle de la psychologie sociale

Comme expliqué plus tôt dans ce travail, les contextes évaluatifs peuvent dans certaines conditions engendrer un sentiment d'infériorité lié à l'identité personnelle ou sociale des individus (Toczek & Souchal, 2017). Le sentiment de menace dû à l'appartenance à un groupe met en lumière un concept particulier, celui des stéréotypes. Selon Dutrévis (2015), les mauvaises réputations qui pèsent sur un groupe et les stéréotypes qui y sont associés participent à l'accroissement de la stigmatisation et des situations discriminatoires vécues par les individus de ce groupe. Goffman résume cette idée en décrivant la notion de stigmatisme comme étant « une trace indélébile qui révèle une dégradation ou une infériorité chez celui qui la possède » (Goffman, 1963, cité par Crahay, 2019, p. 233). Il est cependant important de noter que ce phénomène de stigmatisation est, selon plusieurs auteurs, toujours contextualisé (Dutrévis, 2015 ; Dardenne, 2022). Cela signifie que le stigmatisme peut être vu différemment en fonction du contexte, il peut être considéré comme positif dans une situation et négatif dans une autre.

2.3.2. La menace du stéréotype en situation évaluative

La théorie de la menace du stéréotype (Steele, 1995) offre une explication aux diverses études ayant découvert une différence de performances entre certains groupes dans des contextes évaluatifs. Il a été démontré que les Noirs américains subissaient une pression lorsqu'un stéréotype négatif était activé à leur encontre. La conclusion de cette étude fut que la menace du stéréotype consiste à craindre de confirmer le regard que l'on croit que les autres portent sur son groupe d'appartenance (Steele, 1997, cité par Lafontaine & Toczek, 2023). Autrement dit, la menace proviendrait des croyances qu'ont les individus sur le regard négatif (réel ou fictif) que les autres auraient sur son groupe (Fresson & Dardenne, 2019). Selon Appel et Kronberger (2012) la menace du stéréotype a comme conséquence d'empêcher les individus menacés de développer leurs capacités et de donner le meilleur d'eux-mêmes et donc à être moins performants dans les situations évaluatives. Cependant, Fresson et Dardenne (2019) observent que la menace du stéréotype est un phénomène situationnel. Ils expliquent que c'est le contexte, l'environnement, la situation qui va faire émerger le stéréotype sur un groupe.

Il est intéressant de comparer les différentes procédures d'activation de la menace du stéréotype des divers chercheurs. Comme expliqué ci-dessus, Steele et Aronson (1995) ont par exemple demandé aux sujets de mentionner leur groupe d'appartenance en signalant leur « race ». Dans une autre de leurs études, Steele et Aronson ont soumis un questionnaire aux

participants. D'autres auteurs activent de manière très explicite et saillante le stéréotype. Certains expliquent oralement aux sujets que le but du test est d'observer les différences de performances entre des personnes victimes d'un choc neurologique par rapport aux autres (Fresson, Dardenne & Meulemans 2018, cités par Crahay, 2019). Ils mentionnent donc explicitement l'appartenance de certains sujets à un groupe stigmatisé. Un autre type d'activation de la menace peut se présenter dans la consigne donnée. En effet, la menace peut varier selon la présentation de la tâche, le type ou la discipline de l'évaluation (Goudeau, 2023). Il a par exemple été démontré que les filles étaient moins performantes lorsque l'activité était présentée comme de la géométrie plutôt que du dessin (Huguet, Brunot & Monteil, 2001, cités par Toczek & Souchal, 2017). Le fait que le test soit présenté comme une activité de mathématique, les filles vont alors subir le stéréotype selon lequel elles seraient moins performantes que les garçons. Nous observons donc différentes façons de rappeler les stéréotypes et ce de manière plus ou moins saillante et explicite. La littérature s'est d'ailleurs penchée sur les différences d'effets selon les indices situationnels utilisés. Il apparaît que les indices subtils et les insinuations auraient davantage de conséquences sur les émotions négatives et les pensées intrusives et donc sur les performances (Wheeler & Petty, 2001, cités par Crahay, 2019). Fresson et Dardenne (2019) estiment que les indices situationnels flagrants provoqueraient un rejet plus aisé du stéréotype.

2.3.3. *Les redoublants face aux stéréotypes*

Un élément central semble devoir être pris en compte en ce qui concerne la menace du stéréotype et sur lequel la recherche sur le sujet semble unanime. Pour que la menace s'applique, les individus doivent avoir conscience du stéréotype. Celui-ci doit être connu pour qu'il puisse exercer une influence sur les sujets (Devine, 1989, citée par Crahay, 2019). Cependant, une nuance est à apporter : avoir conscience ne signifie pas adhérer aux stéréotypes. En effet, une personne peut subir la menace du stéréotype sans adhérer à celui-ci (Dutrévis, 2015). Néanmoins, cette découverte selon laquelle les sujets doivent avoir conscience des stéréotypes pour pouvoir en subir les conséquences nous pose question quant à l'âge auquel les enfants sont conscients de ces stéréotypes socialement partagés. Les études menées ont conclu que dès l'entrée en primaire, les enfants pouvaient être conscients de stéréotypes concernant certains groupes (ex: stéréotypes liés au genre, aux origines ethniques, aux statuts économiques) (Fresson & Dardenne, 2019).

Cette recherche s'intéressant au groupe des redoublants, nous nous sommes informés sur la question de la connaissance des stéréotypes liés à ce groupe d'élèves. Peu d'études semblent

s'être penchées sur les stéréotypes visant les redoublants (Dutrévis & Crahay, 2013). Cependant il semble, comme pour les stéréotypes cités ci-dessus, que les enfants développent dès le début des primaires des stéréotypes concernant les redoublants (Fresson & Dardenne, 2019). Crahay (2019), en se basant sur l'étude de Crisafulli et al. (2002), va dans le même sens. Cette étude s'est intéressée aux stéréotypes et aux caractéristiques que les enfants associaient aux enfants redoublants. L'étude a permis de conclure que les caractéristiques attribuées aux élèves redoublants étaient majoritairement négatives (ex : « endormi », « bagarreur », « fufou » « manque d'effort », etc.). Il apparaît donc que les redoublants sont effectivement victimes d'une vision stéréotypée.

À partir de ses observations et des recherches menées sur les stéréotypes chez les enfants, Dutrévis (2015) explique que dès l'école primaire, « les élèves peuvent vivre leurs expériences scolaires sous la menace de confirmer les caractéristiques négatives conférées à leur groupe d'appartenance » (p.68). Nous jugeons donc pertinent d'interroger la façon dont les performances des élèves redoublants sont impactées par ces menaces. D'autant plus que ces enfants sont constamment soumis à cette pression lors des évaluations en classe. Les enfants sont donc régulièrement confrontés aux stéréotypes partagés à leur encontre selon lesquels ils seraient moins performants que les non-redoublants. Crahay (2019) explique que « lorsque ses compétences scolaires sont évaluées (via un exercice ou un contrôle), l'enfant [redoublant] craindrait de confirmer que ce stéréotype est vrai » (p.243). Sur base des différentes découvertes énoncées dans cette revue de la littérature sur la menace du stéréotype, nous pourrions émettre l'hypothèse que ces évaluations constantes entraîneraient des déficits cognitifs et donc des performances inférieures chez les élèves redoublants.

2.4. Le redoublement en Fédération Wallonie-Bruxelles :

Cette étude s'intéressant à un public bien particulier que sont les élèves ayant connu le redoublement, il est nécessaire de définir les caractéristiques de ce groupe. Selon Crahay (2019), les enfants en retard scolaire sont « des élèves qui ont redoublé une ou plusieurs années, ceux qui ont entamé leur scolarité en retard parce qu'ils ont été maintenus en maternelle ou dans d'autres structures éducatives, des enfants venus de l'étranger qui [...] ont repris la scolarité à un niveau inférieur à ce que prescrivait leur âge, etc. » (p.28). Selon Baye (2021), le retard correspond au fait d'être une année en retard par rapport à l'année d'étude correspondant à son âge.

Ce travail de recherche se centre sur un système éducatif bien particulier qu'est celui de la Fédération Wallonie Bruxelles (FW-B). Il est donc utile de s'intéresser à l'ampleur et la place du redoublement et du retard scolaire dans ce système. Lafontaine et ses collaborateurs (2019), qui ont analysé les premiers résultats des tests PISA de 2018, font un constat accablant : seul à peine un peu plus d'un élève sur deux est à l'heure à 15 ans en FW-B. Ils soulignent que « cette tendance de la FW-B à recourir massivement au redoublement est exceptionnelle et est loin de représenter une pratique courante » (Lafontaine *et al.*, 2019, p.28). Depuis l'instauration du décret de 2002 sur le pilotage du système éducatif, la FW-B fournit chaque année « Les indicateurs de l'enseignement ». À partir de ces données, plusieurs observations peuvent être émises. Dans un premier temps, le rapport des Indicateurs de l'enseignement 2022 qui analyse l'année scolaire 2020-2021 souligne que « le pourcentage d'élèves à l'heure diminue de manière quasi linéaire dès la troisième maternelle » (p.46) (et ce jusqu'en sixième secondaire). Dès la troisième primaire, la barre des 10% du nombre d'élèves en retard d'un an ou plus est franchie. En dernière année de l'enseignement secondaire, seuls 45,1% des élèves sont à l'heure. Selon Baye (2021), en FW-B, terminer son secondaire en ayant redoublé une fois ou plus est « la norme ». Une des caractéristiques principales de notre enseignement est donc le recours massif au redoublement en comparaison avec les autres pays européens notamment. En effet, d'après les données PISA, la FW-B avait en 2015 le taux de retard le plus élevé parmi tous les pays de l'Organisation de coopération et de développement économiques (OCDE). En 2015, le taux d'élèves en retard d'au moins une année de 15 ans en FW-B était de 46% tandis que la moyenne des pays de l'OCDE était de 12%. Au vu du contexte particulier que notre enseignement connaît par rapport au redoublement, il nous semble pertinent de nous intéresser à cette population de plus près au travers de cette étude.

2.5. Les mathématiques :

2.5.1. Une discipline pas anodine

Comme exposé plus tôt dans cette présentation théorique, la littérature scientifique continue de proposer plusieurs orientations théoriques quant à la recherche d'explications aux différences de performances entre élèves en situation évaluative, mais aussi en situation d'apprentissage. La psychologie sociale a par exemple permis de remettre en cause la solution innéiste qui justifiait les réussites et les échecs scolaires en invoquant presque exclusivement les facteurs intrinsèques des élèves. Divers auteurs, dont Toczek & Souchal (2017), par le biais de la psychologie sociale, ont par exemple montré à quel point le contexte du traitement de la

tâche était déterminant. Elles ont démontré que les productions cognitives ne peuvent être observées en dehors des dynamiques psychosociales en jeu en situation et qu'elles ne sont donc pas réductibles aux seules caractéristiques de l'apprenant.

En outre, il semblerait qu'un autre facteur soit utile à prendre en compte pour justifier les performances des élèves lors d'un apprentissage ; la matière travaillée. Comme expliqué en introduction de cette étude, l'évaluation mise en place dans le cadre de notre expérimentation se concrétise au travers d'un test et d'un questionnaire relevant du domaine des mathématiques. Les mathématiques font l'objet de nombreuses recherches dans le monde de l'éducation et de l'enseignement. Il est intéressant de noter qu'en termes de performances, les élèves belges ne sont pas les cancrs dans cette discipline. En effet, les scores globaux des élèves de la FW-B en culture mathématique au test PISA en 2022 sont assez semblables à la moyenne des pays de l'ensemble des pays de l'OCDE. Effectivement, la moyenne de la FW-B est, bien que non significative statistiquement, légèrement supérieure à celle de l'OCDE (Baye et al., 2023).

Pourtant, au-delà des chiffres, les mathématiques continuent d'être un centre d'intérêt de nombreuses études depuis de plusieurs années. Blouin (1985), à l'époque déjà, dit de cette discipline qu'elle « suscite à un degré remarquable, toute une variété de réactions dysfonctionnelles d'anxiété, de manque de confiance, de conviction d'impuissance, de manques d'efforts soutenus, de démission prématurée, d'évitement de la tâche, quand ce n'est pas de la matière elle-même » (p.67). Adihou (2011) va même plus loin et n'hésite pas à parler de vraie « souffrance » liée à l'enseignement et l'apprentissage des mathématiques. Cette souffrance naissante face à des tâches mathématiques ne trouve pas sa source dans un seul axe, mais dans divers facteurs de nature affective tels que les difficultés d'ordre cognitif, les blocages affectifs ou encore les pressions familiales, sociales et professionnelles (Adihou, 2011). Les mathématiques en tant que discipline scolaire ne doivent donc pas être considérées comme une matière quelconque.

2.5.2. La question du genre en mathématique

Une problématique très présente dans la littérature scientifique concernant l'étude des mathématiques est la question du genre. De très nombreuses études se sont intéressées aux liens qu'entretiennent le sexe et les performances en mathématiques. En effet, depuis de très nombreuses années, le même constat est établi : les filles sont moins performantes que les garçons (Pajares & Miller, 1994). Une nuance est à apporter, car cette différence de performance ne se marque pas réellement jusqu'à la fin des primaires, mais apparaît et s'accroît par la suite

(Pajares & Miller, 1994). Les résultats du test PISA de 2022 mettent une nouvelle fois en lumière ce phénomène. En mathématiques en 2022 en FW-B, les garçons de 15 ans obtenaient un score de neuf points supérieur aux filles (478 contre 469). Cet écart de performance est identique à celui observé pour la moyenne des pays de l'OCDE (Baye et al., 2023). Une autre donnée importante concernant les différences de performances au test PISA de mathématiques de 2022 concerne la différence de performance entre les élèves à l'heure et ceux en retard scolaire. Le constat est sans appel. En mathématiques, les élèves à l'heure obtiennent un score de 511 points alors que les élèves en retard d'un an ont un score de 427 points. La différence avec les élèves en retard de deux ans ou plus se marquent encore plus, car ceux-ci obtiennent un score de 373 points. Autrement dit, un écart de 84 points sépare les élèves à l'heure de ceux en retard d'un an. Cet écart est de 135 points lorsque les élèves à l'heure aux élèves en retard de deux ans ou plus. Cette variable du retard scolaire est celle qui est le plus liée aux performances dans PISA (Baye et al., 2023).

Cette question de la place du sexe en mathématiques n'est donc pas une problématique propre à notre enseignement. Une fois ce constat posé, il nous paraît utile de s'intéresser à ce que dit la littérature scientifique quant aux causes de ce phénomène. Les raisons évoquées dans la littérature pour expliquer cette différence de performance en mathématique peuvent être de nature assez diverses. Peuvent être cités parmi ces explications, les types d'évaluation (orientée vers la performance et menaçante ou orientée vers la maîtrise et non menaçante), le climat de classe, les attentes des enseignants, les aspects motivationnels, les composantes du soi (anxiété, estime de soi, concept de soi, etc.), mais aussi des facteurs socioculturels comme la place des femmes dans les domaines liés aux mathématiques et les menaces du stéréotype qui y sont liées (Jaegers & Lafontaine, 2018 ; Toczek & Souchal, 2017). Nous allons dans la suite de cette partie théorique, nous intéresser à deux variables liées au soi que sont le concept de soi et l'anxiété. Ces deux concepts cognitifs ne peuvent être négligés lorsqu'il est question d'apprentissage académique et notamment en mathématiques. De plus, ce sont ces deux concepts que nous interrogeons au travers du questionnaire contextuel que nous utilisons dans notre étude.

2.6. L'anxiété liée aux mathématiques

2.6.1. Définition et situation en FW-B

Selon Adihou (2011), l'apprentissage des mathématiques est une tâche complexe qui exige des étudiants un effort cognitif important. En outre, face à une tâche mathématique, des

réactions diverses peuvent apparaître selon les individus (Blouin, 1985). Ces réactions peuvent être positives ou négatives. En effet, selon Roditi (2004), « pour certains les mathématiques sont synonymes de perfection, de refuge, de paix et d'ordre ; pour d'autres, elles représentent un danger, un trou noir, une fatalité » (Roditi, 2004 cité par Adihou, 2011, p. 92). Parmi ces réponses, une réaction émotionnelle est régulièrement associée à l'apprentissage des mathématiques : l'anxiété. Cette réaction humaine naturelle survient chez chaque individu confronté à une situation dangereuse, embarrassante ou stressante (Hashempour & Merhad, 2014). Richardson et Suinn (1972), définissent l'anxiété mathématique comme « caractérisée par un sentiment d'appréhension et de tension qui survient dans les situations impliquant le traitement d'informations numériques » (Richardson & Suinn, 1972, cités par Vilette, 2017, p. 515). L'anxiété peut se présenter sous diverses formes, car elle regroupe différentes sous-réactions telles que l'appréhension et la peur des mathématiques, les émotions négatives, les cognitions et les réactions physiologiques (Weissgerber et al., 2022).

Il semble que les mathématiques comme discipline scolaire engendrent une anxiété très répandue chez les étudiants comparativement à d'autres matières (Weissgerber et al., 2022). Notre enseignement ne fait pas exception à la règle. En effet, les résultats des tests PISA de 2022 mettent en exergue un score d'anxiété vis-à-vis des mathématiques particulièrement important de la part des élèves de 15 ans en FW-B (Baye et al., 2023). Cette synthèse des résultats de PISA de 2022 met également en lumière une accentuation de l'anxiété par rapport aux mathématiques des élèves de la FW-B depuis une dizaine d'années et ce particulièrement pour les élèves moyennement performants (Baye et al., 2023).

2.6.2. Les sources de l'anxiété

Une fois ce constat dressé concernant la situation de notre enseignement, il est intéressant de se demander pourquoi certaines situations scolaires rendent une partie non négligeable des élèves aussi anxieux. Selon Hashempour et Merhad (2014), l'anxiété scolaire naît chez les étudiants « lorsqu'ils ressentent une inquiétude intense à propos de l'incidence à venir et de l'incidence précédente, qu'ils se préoccupent trop d'eux-mêmes et qu'ils s'attachent à agir de manière efficace ou qu'ils sont motivés par divers éléments » (p. 115). La littérature scientifique décèle plusieurs éléments pouvant rendre les situations scolaires, plus précisément celles de mathématiques, particulièrement anxiogènes. Baye et ses collaborateurs (2023) pointent, entre autres, du doigt la difficulté que représente cette discipline pour certains élèves, la peur de l'échec, mais aussi le poids conséquent que cette matière peut avoir sur le parcours scolaire futur des étudiants. Khanna (2015) rajoute que « les tests notés » peuvent eux aussi être source

d'anxiété (Khanna, 2015, cité par Weissgerber et al., 2022). Cette forme d'anxiété est une anxiété scolaire qui n'est plus liée aux mathématiques, mais aux tests. Selon Cavioli et ses collaborateurs (2021), l'anxiété liée aux tests peut être considérée comme « englobant les réactions affectives, cognitives et physiologiques habituellement ressenties lors d'examens ou de situations évaluatives similaires » (Cavioli et al., 2021, p. 364). Ces mêmes chercheurs constatent que ce type d'anxiété se traduit notamment par des cognitions négatives ou des inquiétudes comme des pensées indésirables, incontrôlables et négatives. En outre, l'anxiété liée aux tests peut conduire au fil du temps à une diminution de la motivation et de la capacité à fournir des efforts de la part des élèves pour les tâches scolaires (Hashempour & Merhad, 2014). Bien que l'anxiété liée aux mathématiques et l'anxiété aux tests soient des formes d'anxiété distinctes, elles se distinguent tout de même par des caractéristiques communes, dont la peur de l'échec et certaines réponses cognitives et physiologiques à l'anxiété (Cavioli et al., 2021).

2.6.3. Les conséquences de l'anxiété

L'attention qui est portée par les sciences de l'éducation et la littérature scientifique dans ce domaine à l'anxiété en mathématiques se justifie en grande partie par la relation qu'entretient celle-ci avec les performances des élèves dans cette discipline (Vilette, 2017). En effet, même si, comme vu ci-dessus, l'anxiété liée aux mathématiques et l'anxiété liée aux tests sont bien présentes dans notre enseignement, la question reste de savoir si elles sont forcément néfastes. Selon Weissgerber et ses collaborateurs (2022), il est clair que l'anxiété en mathématiques peut entraîner des effets négatifs sur les apprentissages. Ils soulignent que de manière générale, dans un contexte scolaire anxiogène et hostile, il est difficile de bien apprendre. Ces constats semblent se traduire dans les chiffres. En effet, les résultats PISA de 2022 appuient le fait que l'anxiété est contreproductive. Ces données montrent que les étudiants qui se disent plus anxieux obtiennent un score moindre au test que ceux se disant moins anxieux. Cependant, même s'il semble clair qu'une interaction existe entre l'anxiété et la performance, un lien de causalité franc est à rejeter (Weissgerber et al., 2022). Selon Baye et ses collaborateurs (2023), « il n'est pas possible de déterminer avec certitude si l'anxiété diminue les performances ou si les élèves sont anxieux parce qu'ils ont de faibles performances » (p.59). Bien que ce lien de causalité soit à mettre de côté, la littérature scientifique sur le sujet semble être d'accord sur le fait que l'anxiété exerce une influence sur la performance intellectuelle (Blouin, 1985 ; Cavioli et al., 2021 ; Vilette, 2017). À ce stade, il nous paraît pertinent de s'intéresser aux processus concrets qui permettent à l'anxiété d'affecter la performance des élèves.

L'anxiété est une réaction naturelle qui peut se déclencher chez de nombreuses personnes, mais tout un chacun ne répond pas de la même manière à cette émotion. Pour certains élèves, face à une situation stressante, il est possible que l'anxiété soit bénéfique si elle est faible et contrôlée (Hashempour & Merhad, 2014). À l'inverse, les enfants qui éprouvent une anxiété élevée et non contrôlée ont tendance à interpréter d'une mauvaise manière ou de surestimer la situation dans laquelle ils sont (Hashempour & Merhad, 2014). Effectivement, dans la majorité des cas, les deux formes d'anxiété liées aux mathématiques et aux tests présents à un niveau trop important tendent à avoir un impact négatif sur la performance (Cavioli et al., 2021). Cependant, plusieurs théories distinctes règnent concernant les mécanismes cognitifs qui expliquent cette relation entre l'anxiété et la performance.

La première hypothèse explicative est fondée sur le principe d'interférence cognitive (Weissgerber et al., 2022). Cette explication suppose que l'anxiété, en situation de test notamment, entraîne une altération de la mémoire de travail et une détérioration des fonctions exécutives (Cavioli et al., 2021). Selon Hashempour et Merhad (2014), un niveau d'anxiété élevé consommerait de manière trop importante les ressources nécessaires de la mémoire de travail, réduirait la capacité cognitive durant la phase d'acquisition, d'encodage et de récupération et empêcherait l'étudiant de performer idéalement. Les élèves dans cette situation emploieraient en effet une quantité d'énergie et d'attention trop importante pour tenter de canaliser leur anxiété. Cette énergie ne pourrait plus être déployée entièrement sur la réalisation de la tâche (Hashempour & Merhad, 2014).

La seconde théorie explicative du lien entre anxiété et performance repose sur un principe de déficit (Weissgerber et al., 2022). Cette hypothèse suppose que l'anxiété et la baisse de performance sont dues à une préparation inappropriée au contenu mathématique et à un déficit de compétences dans cette discipline. Selon cette interprétation, c'est un déficit des traitements numériques qui influencerait la performance et donc engendrerait de l'anxiété (Vilette, 2017). En d'autres termes, l'explication par déficit identifie « les déficits d'aptitude en mathématiques comme la cause de l'anxiété et des baisses de performances » (Weissgerber et al., 2022, p.978). La mémoire de travail agirait de manière secondaire dans cette vision. Enfin, cette théorie du déficit se déploie davantage sur un temps long contrairement à la théorie de l'interférence qui agit de façon plus immédiate sur la performance (Weissgerber et al., 2022).

Ces deux explications, bien que différentes, ne sont pas contradictoires ou inconciliables. Elles peuvent coexister. Selon Weissgerber et ses collaborateurs (2022), l'anxiété liée aux

mathématiques et les compétences scolaires ne sont pas opposées et peuvent former une boucle de rétroaction négative.

2.7. Le concept de soi

2.7.1. Définition et caractéristiques du concept de soi

L'anxiété ne semble pas être le seul concept qui influence la manière dont un élève aborde une tâche scolaire. D'autres concepts avec des mécanismes similaires influencent la confiance dont l'étudiant fait preuve face à une situation d'apprentissage tels que le sentiment d'efficacité personnelle, l'estime de soi ou encore le concept de soi (Pajares & Miller, 1994). Bien que ce soit ce dernier concept auquel nous nous intéressons dans notre étude, nous allons ci-dessous distinguer les éléments qui différencient ces concepts.

Le concept de soi est considéré comme la composante cognitive du soi (Martinot, 1995, citée par Martinot, 2001). Selon Raynor et McFarlin (1986), le concept de soi fait référence « aux perceptions qu'une personne a de ses compétences, de ses valeurs, de ses préférences, de ses buts et de ses aspirations » (Raynor & McFarlin, 1986, cités par Sénéchal, 1997, p.12). Paradis et Vitaro (1992) observent que le concept de soi ne correspond pas seulement à la perception, mais aussi à l'évaluation, au jugement qu'une personne porte sur ses caractéristiques personnelles (Paradis & Vitaro, 1992, cités par Galand & Gregoire, 2000). La littérature scientifique relative au concept de soi semble être unanime sur un point ; il est nécessaire de prendre en compte le caractère multidimensionnel du concept de soi. Autrement dit, toute personne possède un concept de soi global qui est lui-même subdivisé en diverses conceptions de soi (Martinot, 2001). En effet, selon le modèle de Shavelson, Hubner et Stanton (1976), il faut différencier cinq catégories de concept de soi : les concepts de soi généraux, sociaux, émotionnels, physiques et académiques (Shavelson, Hubner & Stanton, 1976, cités par Pajares & Miller, 1994). Ce sont ces derniers qui vont attirer notre attention tout particulièrement. Les concepts de soi académiques ou scolaires sont spécifiques à un cours ou une matière. Effectivement, un élève peut avoir un concept de soi positif en mathématiques, mais un concept de soi plus mitigé en sciences ou en français et inversement. De son côté, Strein (1993) distingue deux aspects au concept de soi scolaire. Ce chercheur met en évidence les aspects descriptifs (« J'aime les math ») et les aspects évaluatifs (« Je suis bon en math ») de la perception de soi scolaire (Strein, 1993, cité par Sénéchal, 1997).

De plus, selon Chapman et Tunmer (1995), en ce qui concerne les domaines académiques, le concept de soi serait composé de trois facteurs : « l'attitude (c'est-à-dire l'intérêt) envers le domaine en question, l'impression d'avoir des difficultés (versus des facilités) dans le domaine, et le sentiment de compétence dans ce domaine » (Chapman & Tunmer, 1995, cités par Galand & Gregoire, 2000, p. 4). Le concept de soi est donc une notion complexe qu'il est nécessaire de ne pas confondre avec d'autres concepts fréquemment associés à celui-ci tels que l'estime de soi ou le sentiment d'auto-efficacité (Galand & Gregoire, 2000).

2.7.2. Distinguer le concept de soi d'autres concepts liés

Une notion liée au concept de soi, mais néanmoins différente est le sentiment d'auto-efficacité ou sentiment d'efficacité personnelle. Ces deux concepts sont régulièrement confondus. Bandura (1986) définit le concept d'efficacité personnelle comme « étant le jugement que porte un individu sur ses capacités à organiser et à exécuter une série d'actions dans le but d'atteindre une performance donnée » (Bandura, 1986, cité par Sénéchal, 1997, p.13). D'après cet auteur, l'individu ne focalise plus sa perception sur ses habiletés ou sa personne comme c'est le cas pour le concept de soi, mais plutôt sur sa capacité à accomplir une action spécifique. Selon Pajares et Miller (1994), le concept de soi diffère du sentiment d'efficacité personnelle, car celle-ci est une évaluation contextuelle de la compétence à accomplir une tâche spécifique. Ils rajoutent que même si le concept de soi est spécifique à une matière, il ne l'est jamais à une tâche bien précise. Selon eux, « la question du concept de soi spécifique à un cours, « Êtes-vous un bon élève en mathématiques ? » fait appel à des processus cognitifs et affectifs différents de ceux de la question d'auto-efficacité, « Pouvez-vous résoudre ce problème spécifique ? » » (Pajares & Miller, 1994, p.194). Enfin, il semble donc que le concept de soi et le sentiment d'efficacité ne soient pas si éloignés, mais représentent deux phénomènes à distinguer en raison de la cible de leurs jugements notamment.

L'autre phénomène régulièrement associé et parfois confondu avec le concept de soi est l'estime de soi. Rosenberg (1979) définit l'estime de soi comme correspondant « à la valeur que les individus s'accordent, s'ils s'aiment ou ne s'aiment pas, s'approuvent ou se désapprouvent » (Rosenberg, 1979, cité par Martinot, 2001, p.484). Bien que ces deux champs du soi soient liés, ils ne sont pas équivalents. En effet, un élève peut éprouver une estime de soi globale positive, mais dans le même temps éprouver une conception de soi scolaire et/ou relative à une matière plutôt faible (Harter, 1986, cité par Martinot, 2001). Une fois ces mises au point réalisées concernant deux notions fréquemment associées au concept de soi, il est

pertinent de s'intéresser à l'impact du concept de soi scolaire sur les apprentissages et plus précisément sur les performances académiques des élèves.

2.7.3. Concept de soi scolaire et performance

Selon Galand et Gregoire (2000), le(s) concept(s) de soi est une variable non négligeable qui intervient dans les apprentissages et les performances scolaires. Bandura (1982) va dans le même sens et affirme qu'il existe un lien fort entre ce qu'une personne peut percevoir d'elle-même dans une matière, dans un cours et sa performance dans celui-ci. Pajares et Miller (1994) ont par exemple montré que le concept de soi en mathématiques est lié à la performance en mathématiques, et ce de manière systématique. Calsyn et Kenny (1977) rajoutent que « la réussite scolaire dépend non seulement des performances passées, mais aussi des conceptions de soi actuelles » (Calsyn & Kenny, 1977, cités par Martinot, 2001, p.485). L'analyse des résultats du dernier test TIMSS (Trends in International Mathematics and Sciences Study) de 2023 peut apporter un éclairage sur le sujet. Bien que ce test soit mené dans des classes de quatrième primaire ce qui nous invite à la prudence, les résultats sont parlants. En effet, à ce test, plus les élèves ont un score de perception de soi élevé et se disent confiants, plus leur performance en mathématiques est élevée (Dupont et al., 2024). Une fois ce constat posé, il est intéressant de chercher à comprendre par quels moyens la conception de soi peut influencer la performance. Une des pistes avancées par un certain nombre d'auteurs est le lien qu'entretient le concept de soi avec la motivation et la persistance de l'effort face à une tâche.

En effet, selon Schunk (1991), « les conceptions de soi de réussite reliées au domaine scolaire peuvent influencer la réussite en agissant sur la motivation » (Schunk, 1991, cité par Martinot, 2001, p.485). Byrne (1984) affirme d'ailleurs qu'il ne serait pas concevable d'aborder le concept de soi sans prendre en compte son lien avec la motivation (Byrne, 1984, cité par Sénéchal, 1997). D'après cet auteur, un élève se sentant compétent, ayant une perception de soi positive aura plus de chances d'avoir une bonne motivation intrinsèque. Par conséquent, il s'investira et persévéra plus durement dans la tâche. À l'inverse, un élève faisant appel à des sources de motivation extrinsèque aura tendance à développer un faible engagement dans la tâche et retirera peu de satisfaction. Cela aura pour conséquence d'accroître une faible conception de soi (Sénéchal, 1997).

Enfin, selon Da Fonseca et ses collaborateurs (2004), les jugements que portent les élèves sur leurs compétences les amènent soit à viser la réussite soit à éviter l'échec. Concrètement, les étudiants ayant un concept de soi élevé par rapport au domaine dans lequel la tâche est à

effectuer auront tendance à poursuivre des buts de maîtrise. Ils pourraient également adopter des buts de performance-approche afin de valoriser leur sentiment de supériorité. À l'inverse, les élèves ayant une faible perception de leur compétence adopteront plus facilement des buts de performance-évitement. En conclusion, face à un échec ou une réussite scolaire, il est crucial de ne pas banaliser l'influence que la conception de soi peut avoir ainsi que le lien qu'elle entretient avec la motivation et les buts d'accomplissements poursuivis par les élèves. Outre le fait de prendre en compte cet élément fondamental, il serait pertinent de construire des dispositifs et mettre en place des actions afin de lutter contre les éventuelles influences négatives du concept de soi des enfants.

2.7.4. Focus sur deux catégories d'élèves

Il semble que tous les individus ne considèrent pas de la même manière le concept de soi. L'effet du genre sur la façon dont les personnes perçoivent leurs compétences académiques a par exemple été étudié à plusieurs reprises. Comme nous l'avons vu précédemment, le concept de soi est multidimensionnel (Galand & Gregoire, 2000). Il existe plusieurs types de concept de soi qui peuvent ne pas être équivalents. Sénéchal (1997) a établi plusieurs constats concernant les différences de concept de soi entre les filles et les garçons. Il apparaît que :

- Le concept de soi global des garçons est supérieur à celui des filles.
- Les filles éprouvent un concept de soi scolaire plus élevé que les garçons.
- Les garçons disposent d'un concept de soi supérieur aux filles dans les domaines des mathématiques, de l'activité physique, de l'apparence physique et de la stabilité émotionnelle.
- Les filles présentent un concept de soi supérieur aux garçons concernant la capacité de relation au sexe opposé, les habiletés d'expression verbale ou encore de fiabilité.

Les disparités de concept de soi n'apparaissent pas qu'à la lumière du genre, mais aussi entre les élèves faibles et les élèves forts. Une première donnée importante exprimée par divers auteurs est que de manière générale, l'ensemble des élèves ont tendance à éprouver davantage de conceptions de soi positives que négatives (Chambres & Martinot, 1999 ; Martinot & Monteil, 1995, cités par Martinot, 2001). Cependant, une nuance est apportée, car « les expériences scolaires quotidiennes, notamment chez les élèves en difficulté, peuvent se révéler menaçantes pour le maintien de ces conceptions de soi positives et de l'estime de soi qui leur est liée » (Martinot, 2001, p.492). En effet, les situations entraînant des comparaisons entre élèves, notamment lors des situations d'évaluation, peuvent être néfastes pour les conceptions

de soi des élèves en difficulté, car elles les placent en position d'infériorité. Martinot (2001) explique que face à ce type de situations, les élèves plus faibles vont avoir tendance à se comparer avec les autres élèves en difficulté et non avec les élèves plus forts afin de préserver leur conception de soi. Néanmoins, en faisant cela, les élèves de faible niveau rejettent par la même occasion le principe de réussite scolaire et rendent l'échec commun. Ce phénomène fait référence au principe de « désidentification » qui consiste à diminuer significativement l'importance accordée aux domaines dans lesquels l'individu se juge peu compétent (Crocker et al., 1998, cité par Martinot, 2001, p. 495). Cela permet à l'individu de préserver son estime de soi, mais entraîne une diminution de la quantité de travail ainsi qu'une baisse de la motivation intrinsèque vis-à-vis de ce domaine. Par conséquent, les performances dans ce domaine régressent, l'individu se désidentifie encore davantage pour préserver son estime de soi. Il s'agit d'un véritable cercle vicieux.

2.8. Conclusion de la revue de la littérature

La revue de la littérature nous a permis de mettre en évidence les divers facteurs qui entrent en jeu en situation évaluative. Ces facteurs permettent de déconstruire la conception innée de la réussite scolaire. En effet, les choix posés par les enseignants concernant les évaluations et le contexte qui les entoure ne sont pas sans conséquence. Divers facteurs comme le type d'évaluation, la matière évaluée, les consignes données, le climat de classe, la présence d'autrui influencent leurs performances et impactent les élèves (Lafontaine & Toczek, 2023 ; Huguet, 2011). Ceux-ci ne réagissent pas tous de la même manière à ces diverses variables issues du contexte évaluatif.

Les éléments relatifs au contexte évaluatif ne sont pas les seuls facteurs influençant la performance des élèves. Dans ces situations évaluatives, d'autres variables plus personnelles et relatives au soi des individus entrent également en jeu. La littérature scientifique identifie par exemple l'importance des buts d'accomplissement (influencés par le type de tâche, les objectifs de l'évaluation, la conception de l'intelligence et de la compétence de l'élève, etc.), l'influence et la représentation du soi, le sentiment d'infériorité ou encore les dangers liés aux stéréotypes concernant certaines catégories d'élèves. Les impacts du contexte évaluatif ont été étudiés sur plusieurs catégories d'individus selon le genre, l'ethnie, le statut socio-économique, mais assez peu sur les doublants (Dutrévis & Crahay, 2013 ; Crahay, 2019). Or, selon les Indicateurs de l'enseignement de 2022, notre enseignement possède un taux d'élèves en retard

particulièrement élevé. En outre, il semble que les élèves doublants soient victimes d'une vision stéréotypée et dont ils sont conscients (Fresson & Dardenne, 2019 ; Crisafulli et al., 2002, cités par Crahay, 2019). Les élèves en retard sont régulièrement confrontés aux stéréotypes partagés à leur rencontre selon lesquels ils seraient moins performants que les non-redoublants. Nous jugeons donc pertinent d'interroger la façon dont les performances des élèves doublants sont impactées par ces menaces et sensibles aux variables en jeux en contexte évaluatif.

3. Question de recherche et hypothèses

Cette étude a pour but d'observer les effets du contexte évaluatif d'une évaluation sommative à enjeu en mathématique sur les performances des élèves en retard en troisième secondaire. L'idée est de manipuler le contexte évaluatif de deux manières différentes comme le font Souchal et ses collègues (2014) dans leur étude afin de voir si certaines manières de présenter une évaluation sommative à enjeu peuvent être davantage bénéfiques ou préjudiciables pour les élèves en retard scolaire. Ce mémoire tente donc de répondre à la question suivante :

« Dans quelle mesure les objectifs d'une évaluation sommative avec enjeu peuvent-ils influencer les performances des élèves en retard selon qu'elle soit axée sur la performance (c'est-à-dire une évaluation ayant pour but de principalement comparer et sélectionner les élèves) ou qu'elle soit axée sur la maîtrise (c'est-à-dire une évaluation présentée comme ayant également un but d'apprentissage) ? ».

Pour apporter des éléments de réponse à cette question générale, nous avons amené et présenté cette évaluation de mathématique dans deux conditions différentes. Dans la condition contrôle, cette évaluation est tournée vers des objectifs de performances où des aspects normatifs sont mis en avant au travers de la présentation de cette évaluation (cf. *Tableau 4. Consignes données par l'enseignant selon la condition du dispositif*). C'est cette façon de présenter l'évaluation que nous avons choisi de qualifier de « contrôle » car c'est celle qui se rapproche le plus des évaluations habituellement mises en place dans les classes. Dans la condition expérimentale, nous testons une autre façon de présenter les évaluations sommatives. Cette manière d'utiliser une évaluation sommative tout en n'excluant pas les aspects formatifs, en soutenant l'apprentissage et en mettant l'accent sur des objectifs de maîtrise est dès lors notre condition expérimentale.

Cette façon de procéder nous permettra de comparer les différents résultats au test de mathématique de différentes catégories d'élèves (doublants, non-doublants, filles, garçons, etc.) dans les deux conditions. En outre, notre dispositif comprend également un questionnaire contextuel soumis aux élèves qui interroge trois concepts : les buts d'accomplissement, l'anxiété en mathématique et le concept de soi en mathématique. Observer les différentes réponses des élèves aux trois variables évoquées ci-dessus et les lier aux performances des élèves à l'évaluation de mathématique nous permettra d'apporter une perspective supplémentaire à la façon dont un enseignant peut concevoir ses contextes évaluatifs.

En réponse aux éclaircissements des différents concepts abordés dans la revue de la littérature, nous émettons cette hypothèse principale (*Hypothèse 1a*) : **l'écart de performance entre les deux conditions (contrôle et expérimentale) sera davantage marqué pour les élèves en retard que pour les élèves à l'heure**. À noter que la condition contrôle correspond à la situation où l'évaluation est présentée comme sommative/normative orientée vers la performance et où nous activons par divers moyens le stéréotype relatif aux élèves doublants.

Nous émettons une seconde hypothèse en lien avec la première évoquée ci-dessus évoquant les écarts de performances possibles entre les élèves à l'heure et ceux en retard, car comme nous l'avons déjà précisé, c'est notamment via ce focus sur ce public que se distingue notre étude. Il nous semble également pertinent d'évoquer ici, les liens éventuels à ne pas délaissier entre les variables « sexe » et « retard scolaire », car liées, elles pourraient également apporter des éclairages complémentaires. Nous évoquons d'ailleurs ces pistes au travers de nos hypothèses 1c et 1d.

Hypothèse 1b : l'écart de performance entre les doublants et les non-doublants sera davantage marqué dans la condition contrôle que dans la condition expérimentale.

Nous émettons ces deux premières hypothèses sur base de plusieurs éléments issus de la littérature. Tout d'abord, les évaluations sommatives normatives favorisent la poursuite de buts de performance évitement qui engendrent une baisse des performances chez les élèves en difficulté (Souchal & Toczek, 2010). De plus, comme expliqué dans notre revue de la littérature, l'aspect normatif d'une évaluation peut rappeler des épisodes douloureux vécus par les élèves ayant connu le redoublement ce qui provoquerait chez eux des émotions négatives qui nuiraient à leur performance (Lafontaine & Toczek, 2023). Ensuite, la présence des autres élèves lors d'une évaluation amène les élèves en difficulté à se comparer aux autres et à douter de leurs capacités (Goudeau, 2023 ; Toczek & Souchal, 2017 ; Muller & Butera, 2007). Enfin,

en voulant prouver que le stéréotype les visant est faux, une grosse partie de l'attention des élèves visés est perturbée par des émotions négatives et des pensées intrusives, ce qui fait chuter leur performance (Fresson & Dardenne, 2019). C'est donc sur base de ces éléments extraits de la littérature que nous établissons nos deux hypothèses principales évoquées ci-dessus.

Hypothèse 1c : l'écart de performance entre les deux conditions (contrôle et expérimentale) sera davantage marqué pour les filles que pour les garçons.

Hypothèse 1d : l'écart de performance entre les filles et les garçons sera davantage marqué dans la condition contrôle que dans la condition expérimentale.

Nous émettons ces deux hypothèses sur base des diverses recherches qui ont été menées depuis de nombreuses années concernant la question du genre en mathématiques. Le constat ne semble pas vraiment évoluer au fil du temps. En effet, déjà en 1994 notamment, Pajares et Miller démontraient que les filles obtenaient des résultats inférieurs à ceux des garçons en mathématiques, du moins dès le début de l'enseignement secondaire. Les résultats du test PISA de 2022 mettent une nouvelle fois en lumière ce phénomène : les filles restent moins performantes que les garçons en mathématique. Diverses raisons sont évoquées dans la littérature scientifique pour expliquer cet écart qui perdure (*cf. 2.5.2. La question du genre en mathématique*). Ensuite, nous pensons que les filles obtiendront de meilleures performances dans la condition expérimentale et que l'écart avec les garçons sera plus faible. Notre hypothèse se base, entre autres, sur les travaux menés par Souchal et Toczek (2017). Elles ont démontré que les filles obtenaient de meilleurs scores lorsque l'évaluation notée qu'elles passaient était orientée vers la maîtrise, lorsqu'elle était présentée comme une aide dans leur apprentissage. En orientant l'évaluation vers des enjeux de maîtrise, les chercheuses insistent sur les aspects bénéfiques sur le plan didactique et en mentionnent les bienfaits que les élèves eux-mêmes pourraient tirer de ce test.

L'évaluation de mathématique n'est pas le seul instrument de mesure mis en place dans le cadre du dispositif de cette étude. En effet, un questionnaire contextuel interrogeant trois concepts est soumis aux élèves à la suite du test de mathématique. Ces trois concepts sont les buts d'accomplissement en mathématique, l'anxiété en mathématique et le concept de soi en mathématique. Ce questionnaire complémentaire nous permet d'émettre d'autres hypothèses secondaires concernant le groupe des doublants et le groupe des filles.

Hypothèses 2 : dans la suite de notre questionnaire contextuel, nous demandons aux élèves participants de répondre à six items relatifs à l'anxiété liée aux mathématiques. Par rapport à ce concept, voici les hypothèses que nous émettons :

Hypothèse 2a : **L'ensemble des élèves ayant pris part à la condition contrôle (orientée vers la performance) seront davantage anxieux vis-à-vis des mathématiques que les élèves réunis dans la condition expérimentale (orientée vers la maîtrise).**

Hypothèse 2b : **Les élèves en retard scolaire ayant pris part à la condition contrôle seront davantage anxieux vis-à-vis des mathématiques que les élèves réunis dans la condition expérimentale.**

Hypothèse 2c : **Les filles ayant pris part à la condition contrôle se diront davantage anxieuses vis-à-vis des mathématiques que les filles réunies dans la condition expérimentale.**

Ces trois hypothèses relatives à l'anxiété en mathématiques concernant le retard scolaire et le genre trouvent des sources communes dans la littérature scientifique. En effet, les sources d'anxiété des filles et des élèves doublants peuvent apparaître de facteurs communs. La recherche montre que l'anxiété peut naître face aux mathématiques en tant que discipline scolaire, mais aussi face aux tests. Cette anxiété liée aux tests est ressentie lors d'examens ou de situations évaluatives menaçantes (Cavioli et al., 2021). C'est le cas de notre situation contrôle qui, orientée vers la performance et vers des aspects normatifs, constitue une forme de pression supplémentaire. Or, comme l'ont démontré Souchal et Toczek (2017), les filles sont plus performantes lorsque le contexte évaluatif soutient un climat de maîtrise et non de performance. Cette source d'anxiété additionnée aux stéréotypes socioculturels connus chez les femmes dans le domaine des mathématiques peut rendre les filles en condition contrôle particulièrement anxieuses. Les doublants et les filles qui subissent des stéréotypes négatifs à leur égard entretiennent la pensée selon laquelle ils subiraient un déficit de compétences dans cette discipline pour les filles et dans les tâches scolaire de manière générale pour les doublants (Crisafulli et al, 2007, cités par Crahay, 2019 ; Steele, 1997 ; Souchal & Toczek, 2017). La condition contrôle, où il est dit explicitement que les performances des élèves vont être comparées entre elles, fait écho à une autre dimension que nous souhaitons prendre en compte au vu de notre revue de la littérature : la place de la comparaison sociale en situation évaluative. Face à cette annonce, les doublants et les filles vont imaginer que les autres sont plus

performants qu'eux, ce qui va engendrer toute une série de pensées intrusives, du stress ainsi que de l'anxiété (Goudeau, 2023 ; Lafontaine & Toczec, 2023 ; Muller & Butera, 2007).

Hypothèses 3 : dans la suite de notre questionnaire contextuel, nous demandons aux élèves participants de répondre à trois items relatifs au concept de soi en mathématique. Par rapport à ce concept, voici les hypothèses que nous émettons :

***Hypothèse 3a* : l'ensemble des élèves ayant pris part à la condition contrôle (orientée vers la performance) auront une conception de soi liée aux mathématiques plus faible que les élèves réunis dans la condition expérimentale (orientée vers la maîtrise).**

***Hypothèse 3b* : les élèves en retard scolaire ayant pris part à la condition contrôle exprimeront une conception de soi liée aux mathématiques plus faible que les élèves en retard réunis dans la condition expérimentale.**

***Hypothèse 3c* : les filles ayant pris part à la condition contrôle exprimeront une conception de soi liée aux mathématiques plus faible que les filles réunies dans la condition expérimentale.**

Ces trois hypothèses relatives au concept de soi en mathématiques concernant le retard scolaire et le genre trouvent des sources communes dans la littérature scientifique. En effet, les diverses études s'intéressant au concept de soi ont permis de démontrer que les situations scolaires menaçantes entraînant des comparaisons entre élèves (comme c'est le cas dans notre condition contrôle) peuvent dégrader le concept de soi des élèves en difficulté (Martinot, 2001). Les doublants font potentiellement partie de cette catégorie d'élèves. Cette situation de comparaison avec autrui place ces élèves en situation d'infériorité (Martinot, 2001). Concernant les filles, la recherche montre qu'elles ont un concept de soi scolaire plus élevé, mais un concept de soi plus faible dans le domaine des mathématiques par rapport aux garçons (Sénéchal, 1997). Il sera intéressant de voir si le concept de soi diffère en condition contrôle et en condition expérimentale, car comme nous l'avons expliqué précédemment, en situation d'évaluation orientée vers la performance, les élèves développent davantage des buts de performance, ce qui ne favorise pas la motivation et la persévérance (Darnon & Butera, 2005). Cela aura pour conséquence d'accroître une faible conception de soi selon Sénéchal (1997).

Hypothèse 4 : dans notre questionnaire contextuel, nous demandons aux élèves participants de répondre à douze items relatifs aux buts d'accomplissement. Par rapport à ce concept, voici les hypothèses que nous émettons :

Hypothèse 4a : **La condition contrôle, portée vers la performance, devrait pousser davantage l'ensemble des élèves à adopter des buts de performances comparativement à la condition expérimentale axée vers la maîtrise.**

Hypothèse 4b : **Les élèves en retard poursuivront moins de buts de maîtrise-approche dans la condition contrôle orientée vers la performance que les élèves en retard dans la condition expérimentale orientée vers la maîtrise.** La revue de la littérature nous a permis de montrer que le type d'évaluation influençait la motivation, l'intérêt des élèves ainsi que la représentation que se font les élèves de la tâche (Butera, 2023). Par conséquent, cela impacterait leurs performances. En effet, les évaluations aux orientations normatives entraînent les élèves à adopter des buts d'accomplissement dits de performances qui poussent les élèves à vouloir être meilleurs que les autres (Souchal et al., 2014).

Hypothèse 4c : **Les élèves en retard poursuivront davantage de buts de performance-évitement dans la condition contrôle orientée vers la performance que les élèves en retard dans la condition expérimentale orientée vers la maîtrise.** Selon la littérature scientifique, lorsqu'ils se sentent menacés ou en difficulté, les élèves vont avoir tendance à douter d'eux-mêmes, à se décourager et à adopter des stratégies peu efficaces (Ames, 1984 ; Dweck, 1999, cités par Souchal & Toczek, 2010). Ces élèves vont avoir tendance à adopter des buts de performance-évitement, c'est-à-dire éviter d'échouer, ne pas faire moins bien que les autres, éviter de paraître incompetents (Darnon & Butera, 2005). Contrairement aux buts de performance-approche qui peuvent aider certaines catégories d'élèves comme les garçons à obtenir de meilleures performances (Duda, 1988, cité par Toczek & Souchal, 2017), poursuivre des buts de performance-évitement entraînerait presque systématiquement des effets délétères sur les performances (Souchal & Toczek, 2010).

4. Méthodologie

4.1. Méthode choisie

Pour ce mémoire, nous planifions de faire appel à une méthode quantitative. Selon les auteurs, ce type de méthode peut aussi être qualifié de nomothétique (Paquay, 2006) ou confirmatoire (Schwab & Held, 2020). Ces différentes dénominations représentent néanmoins toutes la même méthode. Cette méthode quantitative vise à expliquer ou prédire un phénomène, ainsi de confirmer ou infirmer des hypothèses émises préalablement (Schwab & Held, 2020).

C'est ce que nous tenterons de réaliser dans ce mémoire. Nous établirons un dispositif qui nous permettra de confirmer ou d'infirmer nos hypothèses (cf. 3. *Question de recherche et hypothèses*). Nous avons donc, à partir de notre question de recherche, choisi ce type de méthode, car nous souhaitons en effet vérifier si les situations évaluatives que nous mettrons en place impacteront les performances des élèves doublants comme nous l'imaginons.

La méthode quantitative présente généralement un plan d'investigation ciblé sur des groupes afin de recueillir des données standardisées et effectuer une analyse statistique de ces données (Paquay, 2006). C'est pourquoi nous avons établi une étude quasi expérimentale. Ce type d'étude permet de comparer deux groupes : un groupe expérimental et un groupe contrôle (Slavin, 2007). Nous comparons effectivement la performance de différentes catégories d'élèves (doublants/non-doublants, filles/garçons, etc.) en condition contrôle, en situation évaluative sommative et normative forte à la performance de ces mêmes groupes d'élèves en condition expérimentale, c'est-à-dire en situation d'évaluation sommative aux aspects formatifs. Les performances au test de mathématique ne seront pas les seules données que nous analysons, car nous comparons également les réponses des élèves des deux conditions au questionnaire contextuel portant sur les buts d'accomplissement ainsi que l'anxiété et le concept de soi en mathématique.

4.2. Description de l'échantillon

Quatre cent quatre-vingt-cinq élèves de troisième secondaire issus de l'enseignement général et de l'enseignement technique de transition en FW-B ont participé à cette étude. Cet échantillon global a été réparti en deux groupes ; soit dans la condition contrôle (CC), soit dans la condition expérimentale (CE). La répartition des filles et garçons est relativement équilibrée. C'est également le cas concernant le retard scolaire puisque seul un écart de trois élèves sépare le nombre d'enfants doublants entre les deux conditions (cf. *Tableau 1 : répartition des élèves de l'échantillon dans les deux conditions du dispositif en fonction du genre et du retard scolaire*). En ce qui concerne les taux de doublants dans nos échantillons, nous pouvons noter qu'ils sont assez équivalents aux taux mesurés par PISA dans l'ensemble de la FW-B. En effet, selon les Indicateurs de l'enseignement 2022, dans l'enseignement général, 22% des élèves étaient en retard. Dans l'enseignement technique de transition ce taux est de 44%. Le tableau ci-dessous détaille les compositions des deux échantillons selon le genre et le retard scolaire :

Tableau 1 : répartition des élèves de l'échantillon dans les deux conditions du dispositif en fonction du genre et du retard scolaire

	Condition contrôle	Condition expérimentale
Total	245	240
Doublants	39	42
Non-doublants	206	198
Filles	108	113
Garçons	137	127

Chaque condition comprend douze classes au total. Comme expliqué ci-dessus, ces classes sont issues de l'enseignement général (GT) ou de l'enseignement de technique de transition (TT). Ces classes ont été sélectionnées sur base volontaire de la part de leur professeur de mathématiques. Ces enseignants ont été approchés par mail ou téléphone via les directions d'école, les secrétariats ou des connaissances du chercheur. Au total, quatorze professeurs issus de douze écoles de la FW-B ont accepté de prendre part à notre expérimentation. Dans notre demande initiale de participation à cette étude, il était imposé aux enseignants intéressés d'avoir plus d'une classe de mathématiques en troisième secondaire. Ce système nous permettait de placer aléatoirement une classe A de l'enseignant dans la CC et une classe B de ce même enseignant dans la CE. Cela permettait d'avoir une équivalence d'échantillon dans les deux conditions puisque bien que les élèves et les contextes de classe soient différents, les deux classes appartiennent à la même école, ont le même enseignant et donc, toute proportion gardée, la même manière d'enseigner, la même façon de proposer des évaluations, etc. Cependant, la réalité des faits, la difficulté de trouver des enseignants et des classes disposés à participer à notre étude nous a forcé à revoir ce critère de minimum deux classes par enseignant. C'est pourquoi il faut noter que sur les douze classes, deux enseignants n'ont qu'une seule classe participante, et ce dans les deux conditions. Autrement dit, quatre enseignants ayant pris part à notre étude n'ont eu qu'une classe dans une des deux conditions. Tous les autres professeurs ont au moins deux classes qui ont été réparties aléatoirement dans les deux conditions.

Une des problématiques principales de la mise en place de notre étude sur le terrain fut d'obtenir deux échantillons équivalents de manière à pouvoir comparer toute une série de données que nous détaillerons plus tard. Même si obtenir deux groupes totalement comparables est impossible, nous avons tout fait pour que ce soit le plus possible le cas. Nous avons donc soumis aux enseignants et aux élèves participants un questionnaire contextuel afin de pouvoir

dessiner une description de nos deux échantillons. Le tableau ci-dessous détaille les critères qui nous permettent de dire que bien qu'ils ne le soient pas à cent pour cent, nos deux échantillons sont comparables.

Tableau 2 : grille de description et de comparaison des échantillons deux conditions

	Condition contrôle	Condition expérimentale
N classes	12	12
N classes GT	11	10
N classes TT	1	2
N enseignants ayant une seule classe	2	2
N classes ayant eu la période de rappel	4	4
N classes ayant vu la matière au Q1	3	2
N classes ayant vu la matière au Q2	9	10
N classes ayant travaillé la matière entre 20 et 25 périodes cours	11	11
N classes rassemblant des Es d'une même option	6	7

Cette grille exposant différents critères nous permet de nous dire que nos deux échantillons sont relativement comparables. Il nous semble également intéressant de notifier deux éléments supplémentaires. Le premier concerne le nombre de périodes de cours par semaine que les élèves suivent en mathématiques. Il est en réalité de cinq pour tous les élèves toutes classes et conditions confondues. En effet, dans l'enseignement de transition, tous les élèves ont cinq périodes de mathématiques par semaine. Cet élément a été un facteur capital dans notre choix de s'adresser à des classes de troisième secondaire, car sans cela, la comparabilité de nos deux échantillons aurait pu être affectée.

Le second élément concerne les options que suivent les élèves. En effet, dans notre enseignement, les élèves de troisième secondaire peuvent déjà décider de poursuivre certaines options de cours comme les sciences économiques, le latin, les sciences sociales ou encore le sport. Dans le questionnaire contextuel rempli par les enseignants de mathématiques, il leur est demandé de spécifier, si c'est le cas, l'option qui rassemble les élèves d'une même classe. Il ressort que six classes sur douze de la CC ne rassemblent pas des élèves d'une option

particulière. Dans la CE, cinq classes sur les douze sont dans ce cas de figure. Malheureusement, il était impossible de pouvoir contrôler cette variable dans notre dispositif. Néanmoins, il nous semblait intéressant de connaître cette statistique et important de le spécifier ici. Les autres classes ne rassemblant pas d'élèves d'une même option regroupent des élèves suivant des options différentes. Le choix de réunir des élèves issus de mêmes options dépend des écoles.

Enfin, il nous semble nécessaire d'explicitier la manière dont les échantillons ont été composés. En effet, l'attribution des classes à la condition expérimentale ou contrôle a été déterminée par le chercheur et non par l'enseignant. Ainsi, comme expliqué ci-dessus, des « paires » de classes ont été constituées de façon à ce qu'une classe de la CC et une classe de la CE aient le même professeur de mathématique. Dans le cas où les classes rassemblaient des élèves d'une même option, nous avons veillé à répartir équitablement les différents types d'options dans les deux conditions. Dans le cas contraire, si les classes de l'enseignant ne rassemblaient pas d'élèves d'une même option, nous avons attribué les classes aux conditions de manière aléatoire. La même démarche a été appliquée avec les quelques enseignants n'ayant qu'une classe. Si la classe rassemblait des élèves d'une même option, le chercheur plaçait celle-ci dans une des conditions en veillant à équilibrer les options dans les deux échantillons, mais si ce n'était pas le cas, celle-ci était positionnée dans une condition de manière aléatoire. En appliquant cette procédure, nous souhaitons éviter que certains enseignants et/ou le chercheur soient tentés de choisir leurs « meilleures classes » pour telle ou telle condition. En outre, en procédant de cette manière, nous souhaitons obtenir au maximum une équivalence de niveaux des élèves dans les deux conditions.

4.3. Procédure générale

Le tableau ci-dessous représente les différentes étapes du dispositif de notre étude mis en place dans les classes participantes. Ce schéma ne mentionne pas les divers échanges pré et post test avec les enseignants ainsi que les diverses étapes administratives liées à cette recherche.

Tableau 3 : *présentation simplifiée des étapes concrètes du dispositif de cette recherche*

Période 1	Période 2
* <u>Période de rappel</u> sur la matière du test proposé en deuxième période (Chapitres UUA3 et UAA4 sur les fonctions).	* <u>Questionnaire contextuel 1</u> (informations personnelles relatives aux élèves) * <u>Test</u> (évaluation de mathématique sur les fonctions)

* <u>Annonce du test</u>	* <u>Questionnaire contextuel 2</u> (sur les buts d'accomplissement, l'anxiété et le concept de soi en mathématique)
--------------------------	--

Le dispositif que nous avons mis en place et auquel les enseignants ont accepté de participer se compose donc de deux périodes de cours. Il est important de noter que ce sont les enseignants, professeurs de mathématiques de leur(s) classe(s), qui dirigent et mènent l'ensemble du dispositif. Le chercheur n'est pas présent en classe. Afin de s'assurer que tout se déroule dans les conditions prévues de l'étude, le chercheur a organisé avec chaque enseignant un entretien ou un appel téléphonique afin d'expliquer celle-ci, clarifier les zones d'ombre et répondre aux questions du professeur. De plus, un guide de passation a été fourni par le chercheur aux enseignants (Annexe 1). Ce guide, inspiré des guides de passation des épreuves externes, détaille de manière précise toutes les étapes par lesquelles l'enseignant doit passer ainsi que toutes les consignes importantes qui doivent absolument être communiquées aux élèves. De cette manière, nous veillons à nouveau à ce qu'il y ait une cohérence et une équivalence de passation entre les classes et entre les deux conditions.

Période 1 : cette première période consiste en un rappel, un cours dispensé par le professeur participant pour réviser la matière du test qui sera soumis aux élèves (cf. 4.4. *Instruments de mesure*). Nous avons au préalable réalisé, en collaboration avec une professeure de mathématique, une fiche sommaire reprenant les différents points du chapitre à revoir avec les élèves (Annexe 2). De cette manière, le contenu du test est rappelé et les élèves ne sont pas interrogés sur un sujet qu'ils ne connaissent pas ou qu'ils ont totalement oublié. Comme mentionné plus tôt (cf. 4.2. *Description de l'échantillon*), environ la moitié des classes des deux conditions n'ont pas réalisé cette heure de rappel, et ce pour deux raisons. Premièrement, les professeurs dans ce cas n'étaient pas en capacité de consacrer deux périodes de cours à notre recherche. Deuxièmement, ces enseignants étaient en réalité en train de travailler la matière évaluée dans notre test. Si et seulement si les enseignants répondaient à ces critères, alors il était accepté qu'ils ne réalisent pas cette première séquence. Pour les professeurs qui ont vu la matière plus tôt dans l'année et qui réalisaient ce rappel, il leur a été demandé de maintenir un temps d'une semaine maximum entre les deux périodes, et ce afin de ne pas créer un décalage temporel trop important.

Période 2 : La récolte de données en tant que telle est prévue lors de cette deuxième période. Elle est organisée de telle sorte qu'elle dure une heure de cours. Comme présenté dans le Tableau 3, elle est répartie en trois étapes. Tout d'abord, les élèves remplissent un premier

questionnaire contextuel qui nous permet de récolter des informations sur nos sujets (nom, prénom, âge, sexe et retard scolaire éventuel). Ensuite, les étudiants répondent au test de mathématiques tiré d'une EENC de 2017 sur les fonctions durant trente minutes (Annexe 3). Ce qui différencie les deux conditions est la façon de présenter l'évaluation. En effet, nous comparons la performance des élèves en situation évaluative axée sur la performance (CC) à celle en situation d'évaluation axée sur la maîtrise (CE). Mis à part cela, les deux groupes passent le même test. Afin de comprendre réellement ce qui différencie les deux conditions, vous trouverez ci-dessous un tableau reprenant la principale consigne donnée par les enseignants pour introduire cette évaluation dans chaque condition. C'est à travers cette consigne que nous rendons l'évaluation axée sur la performance (CC) ou sur la maîtrise (CE). Ces consignes sont inspirées de celles utilisées par Souchal et ses collaborateurs (2014). Comme nous l'avons déjà mentionné précédemment, notre dispositif est en partie inspiré du leur, car elles aussi ont manipulé les consignes données afin de modifier la façon dont l'évaluation de sciences qu'elles mettaient en place était présentée.

Tableau 4 : *Consignes données par les enseignants selon la condition du dispositif*

CC (axée performance)	<i>« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test permettra d'évaluer votre niveau dans cette matière. Ce test nous permettra également de comparer vos capacités à celles des autres étudiants de la classe. »</i>
------------------------------	---

CE (axée maîtrise)	<i>« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test vous aidera à comprendre davantage ce que vous avez vu du chapitre sur les fonctions. Vous verrez que, même pendant le test, vous continuerez à apprendre et à développer des connaissances. »</i>
---------------------------	--

Une fois le temps écoulé, l'enseignant reprend les feuilles et distribue le deuxième questionnaire contextuel (Annexe 4). Celui-ci interroge trois concepts : les buts d'accomplissement, l'anxiété liée aux mathématiques et le concept de soi en mathématiques. Une dernière donnée importante est à préciser quant au déroulement de ce dispositif : les enseignants n'ont pas accès à l'évaluation avant la passation du test. Cela permet d'éviter que certains d'entre eux aiguillent ou orientent de manière trop importante leurs élèves sur le contenu à travailler et à réviser. De plus, cela aurait mis en difficulté l'équivalence entre les échantillons que nous avons essayé de maintenir le plus possible.

4.4. Instruments de mesure

Comme expliqué précédemment, notre dispositif comprend deux outils principaux : un test de mathématique et un questionnaire contextuel. Nous détaillerons ci-dessous les éléments qui nous ont poussés à sélectionner ces instruments ainsi que les avantages et les limites que nous avons décelés les concernant. Nous évoquerons également les changements et les modifications que nous avons apportés à ceux-ci afin qu'ils soient davantage en adéquation avec nos objectifs.

4.4.1. Test de mathématique

Lorsque les enseignants ont donné les consignes relatives à chaque condition (*cf.* 4.3. *Procédure générale*), les élèves reçoivent les feuilles. La première chose à faire est de remplir un court questionnaire contextuel visant à obtenir des informations contextuelles (nom, prénom, âge, sexe, retard scolaire éventuel et option suivie). Ensuite, l'évaluation de mathématique proprement dite commence (*Annexe 3*). Ce test est en réalité une partie de l'Évaluation Externe Non Certificative (EENC) de mathématique de 2017 portant sur les fonctions. Les items de ce test portent sur les chapitres UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré). Ce test permet de répondre à notre question de recherche ainsi que de confirmer ou infirmer nos hypothèses 1a, 1b et 1c.

Afin que le test puisse être réalisé en trente minutes, nous avons sélectionné sept questions. La sélection de ces questions issues de cette EENC s'est effectuée sur base de deux critères. D'une part, les questions choisies devaient porter sur des raisonnements et non sur des points théoriques spécifiques comme des définitions de manière à ne pas mettre les élèves en difficulté face à des points de matière précis. D'autre part, les questions devaient être suffisamment complexes et discriminantes. Cet exercice de sélection a été réalisé en collaboration avec une enseignante de mathématique afin d'assurer une certaine cohérence. Les sept exercices sélectionnés regroupent seize items. Chaque item peut être classé parmi trois processus : Connaitre, Appliquer et Transférer.

Une des limites que nous pouvons identifier dans le fait d'avoir choisi une EENC pour notre étude se situe, selon nous, dans l'esprit de ce type d'évaluation. En effet, comme leur nom l'indique, les épreuves non certificatives poursuivent des objectifs formatifs et ne peuvent sanctionner les élèves dans leur parcours. Dans le cadre de notre étude, nous rendons cette

EENC sommative à enjeu, ce qui détourne légèrement l'objectif premier de ce type d'évaluation.

Néanmoins, notre choix de se porter sur une EENC pour mettre en place ce dispositif est motivé par pléthore d'éléments. Tout d'abord, ce type d'évaluation est développé par un groupe de travail composé de différents acteurs du monde éducatif. Ce groupe choisit les compétences à évaluer dans la discipline et conçoit l'épreuve en prenant en considération toute une série de facteurs organisationnels (validité interne, cohérence des items, clarté des corrections, des consignes ou du vocabulaire employé, etc.). De plus, avant d'être validée, chaque EENC est prétestée afin d'en assurer la pertinence et la validité. Ensuite, ce type d'évaluation externe permet d'assurer un niveau de difficulté adéquat et adapté à l'ensemble des écoles de l'enseignement de transition de la FW-B. En effet, une évaluation réalisée par le chercheur ou par un professeur de mathématique unique n'aurait pas permis d'assurer cette équité pour les classes sélectionnées de notre étude. Les grilles de correction fournies avec les EENC permettent de réduire les biais de subjectivité de la part des correcteurs. Dans le cadre de notre étude, le chercheur est le correcteur. Le système de correction et de codage utilisé par le chercheur est identique à celui utilisé dans le cadre des EENC à savoir ; code 1 pour une réponse correcte, code 0 pour une réponse incorrecte et code 9 pour une absence de réponse. Enfin, un des avantages de ce test de mathématique est qu'il porte sur de la matière de troisième secondaire. Cela permet aux enseignants qui acceptent de participer à notre étude de ne pas perdre de temps de cours pour autre chose. Enfin, cela permet aux professeurs de situer le niveau de leurs élèves par rapport aux exigences d'une évaluation externe.

4.4.2. Questionnaire contextuel

Le deuxième instrument de mesure employé dans notre dispositif est un questionnaire contextuel (Annexe 4). Ce questionnaire interroge trois concepts : les buts d'accomplissement, l'anxiété liée aux mathématiques et le concept de soi en mathématique. Les données obtenues nous permettront de les relier aux performances aux tests et de confirmer ou infirmer nos hypothèses 2, 3 et 4 qui portent sur ces concepts-là. La passation de ce questionnaire s'est déroulée juste après la fin du test de mathématique.

La première partie du questionnaire portant sur les buts d'accomplissement est inspiré de celui de Darnon et Butera (2001). Ces deux auteurs se sont attachés à traduire en français l'échelle des buts d'accomplissement d'Elliot et McGregor (2001). Bien que ce travail de

traduction puisse nous amener à la prudence, les diverses analyses réalisées par Darnon et Butera concernant la validité interne et externe de leur échelle traduite est rassurante. En effet, selon eux, « l'ensemble des résultats d'Elliot et McGregor (2001) est reproduit avec la version française du questionnaire [...] ce résultat est conforme à ce qui est obtenu dans d'autres études (Nolen, 1988 ; AL-Emadi, 2001) » (p. 123). Cette échelle sur les buts d'accomplissement compte douze items, trois pour chaque but d'accomplissement.

L'autre partie de ce questionnaire contextuel porte sur le concept de soi en mathématique et sur l'anxiété liée aux mathématiques. La première notion concerne les trois premiers items tandis que le second concept porte sur les six derniers items de la section 1. Ces échelles sont tirées de l'enquête PISA 2022. Le fait que ces échelles aient été configurées dans le cadre d'une enquête de cette renommée nous assure une qualité certaine. De plus, avant d'être validées, ces échelles ont été prétestées afin d'en assurer la pertinence et la validité. Ensuite, comme dans notre recherche, les tests PISA ont lieu à l'âge de quinze ans, en troisième secondaire. Cela assure un niveau de difficulté des items et du vocabulaire adéquat et adapté à l'ensemble des écoles de l'enseignement de transition de la FW-B.

La présentation méthodologique de cette étude étant réalisée, nous allons par la suite aborder la présentation et l'analyse des données recueillies. Cette partie aura pour objectif d'observer, comparer et ensuite analyser les résultats que nous avons obtenus au travers de nos divers instruments de mesure présentés préalablement. Nous veillerons également à mettre en avant les résultats qui nous permettront de confirmer ou réfuter nos hypothèses.

5. Présentation, analyse et discussion des résultats

5.1. Consistance interne des instruments de mesure

5.1.1. Validation du questionnaire contextuel

Comme abordé succinctement dans la partie méthodologie, les variables employées dans notre questionnaire contextuel (les buts d'accomplissement, le concept de soi en mathématiques et l'anxiété liée aux mathématiques) sont mesurées à l'aide d'une même forme d'échelle : une échelle de Likert. Celle-ci offre les réponses suivantes aux répondants : « pas du tout d'accord », « pas d'accord », « d'accord » et « tout à fait d'accord ». Chaque réponse correspond à un score (de 1 à 4). Ce type d'échelle comme n'importe quelle autre a des

inconvenients, mais aussi des avantages, notamment en ce qui concerne notre étude. Elle ne permet pas aux élèves de sélectionner une réponse neutre comme « je ne sais pas » ou « je préfère ne pas répondre ». Selon Lafontaine (2018), « si on est quasi certain que les sujets ont un avis sur la question, il est parfois préférable d'utiliser des échelles paires (sans point central) pour forcer le choix [...] » (p. 14). Autrement dit, ce type d'échelle force les élèves à prendre position et c'est ce que nous souhaitons, car les items de notre étude ne demandent pas de répondre à une question de connaissance par exemple, mais d'émettre un avis sur ce qu'ils ressentent et éprouvent dans les situations évoquées par nos items. Cette manière de procéder permet en outre d'éviter l'effet de tendance centrale (Lafontaine, 2018). Avant d'aller plus loin dans l'analyse proprement dite des données récoltées au travers de notre questionnaire contextuel, il est primordial d'observer la cohérence interne de notre instrument afin de valider nos échelles de mesure. Pour ce faire, nous nous intéressons ci-dessous aux différents Alphas de Cronbach que nous avons calculés (*cf. Tableau 5 : Alphas de Cronbach des constructs étudiés*). Ces données nous permettent de vérifier la cohérence interne de nos échelles, leur unidimensionnalité. Autrement dit, cela nous permet de voir si nos échelles mesurent bien un seul et même construct (le concept de soi en mathématiques, l'anxiété liée aux mathématiques ou les quatre buts d'accomplissement).

Une des limites de l'Alpha de Cronbach est qu'il dépend de manière importante à l'échantillon dans lequel le questionnaire a été soumis. Selon Navarro et Foxcroft (2020) « un échantillon biaisé, non représentatif ou de petite taille pourrait produire un coefficient alpha très différent de celui d'un grand échantillon représentatif » (p. 479). Or comme nous l'avons détaillé dans la partie méthodologie (*cf. 4.2. Description de l'échantillon*), nous ne prétendons pas avoir dans le cadre de notre étude des échantillons parfaitement représentatifs, car ce sont des échantillons de convenance, mais le nombre important d'élèves ayant participé à cette recherche nous permet d'avoir un nombre de sujets élevé et de contrecarrer cette limite de l'alpha évoquée ci-dessus. Nous pouvons donc considérer que nos coefficients de Cronbach seront plutôt fiables. Le tableau 5 ci-dessous met en évidence les six constructs mesurés au travers de notre questionnaire contextuel en les associant à leur acronyme, le nombre d'items qui s'y rapporte ainsi que leur alpha.

Tableau 5 : Alphas de Cronbach des constructs étudiés dans le questionnaire

Constructs	Acronymes	Nombre d'items	Alpha
Concept de soi lié aux mathématiques	SOI	3	0,60
L'anxiété liée aux mathématiques	ANX	6	0,78
Les différents types de buts :			
Maitrise approche	BMA	3	0,72
Maitrise évitement	BME	3	0,71
Performance approche	BPA	3	0,81
Performance évitement	BPE	3	0,42

Le construct du concept de soi comprend trois items. Il obtient un coefficient de Cronbach de 0,60. Cet indice ne peut pas être qualifié de « bon », mais il peut être considéré comme « suffisant » (Navarro & Foxcroft, 2020). Au vu du faible nombre d'items pour ce construct, nous avons décidé de ne pas en supprimer et de se satisfaire de ce taux. L'indice d'anxiété par rapport aux mathématiques est basé sur six items. Ce construct obtient un alpha de 0,78, ce qui est un bon taux (Navarro & Foxcroft, 2020). En ce qui concerne les alphas de Cronbach des quatre buts d'accomplissement, il peut être conclu que ceux-ci sont dans l'ensemble suffisants. En effet, les constructs relatifs aux buts de maitrise approche et évitement ainsi que ceux relatifs aux buts de performance approche obtiennent des coefficients de Cronbach supérieurs à 0,7. Seul le construct relatif au but de performance évitement n'est pas bon. Il est à noter que dans les quatre buts, supprimer des items n'aurait pas permis d'obtenir des alphas de Cronbach plus élevés. De plus, supprimer un item alors que ces constructs n'en possèdent que trois initialement n'aurait pas tellement de sens. En conclusion, à partir des coefficients alpha calculés pour chaque construct, nous pouvons affirmer que les deux séquences du questionnaire contextuel de notre étude ont une bonne consistance interne (voire acceptable pour le premier). En effet, cinq des six constructs présentent un alpha de Cronbach satisfaisant ou élevé. Ces différents coefficients nous permettent d'affirmer que ces items de nos échelles mesurent bien un seul et même construct.

5.1.2. Validation du test de mathématique

Nous avons également veillé à analyser la cohérence interne de notre évaluation de mathématique. Comme nous l'avons détaillé dans la partie méthodologie (*cf. 4.4.1. Test de mathématique*), ce test est tiré d'une EENC de 2017, ce qui peut lui assurer une fiabilité certaine

au vu des experts impliqués pour sa conception. Cependant, celle-ci étant trop longue pour pouvoir s'intégrer dans notre dispositif, nous l'avons modifiée de manière à ce qu'elle puisse être réalisée en trente minutes par les élèves. Modifier cette évaluation nous fait donc prendre le risque d'éventuellement détériorer sa validité. Il est donc pertinent de vérifier la consistance interne de notre évaluation proposée aux élèves. Le tableau 6 ci-dessous présente les alphas de Cronbach des trois catégories d'items et le coefficient calculé pour l'ensemble du test. Pour rappel, les items sont des questions relatives aux compétences « Connaitre » (C1), « Appliquer » (C2) ou « Transférer » (C3).

Tableau 6 : *Alphas de Cronbach relatifs au test de mathématique*

	Items correspondants	Alpha de Cronbach
Connaitre – C1	Items 1a, 1b, 1c, 1d Items 4a, 4b, 4c	0.62
Appliquer – C2	Items 2a, 2b Item 5 Items 6a, 6b	0.43
Transférer – C3	Items 3a, 3b Items 7a, 7b	0.38
Ensemble du test	Tous	0.70

Le premier constat que nous pouvons poser en analysant ces différents coefficients est que seul l'un d'entre eux atteint le seuil des 0.7 qui permet de qualifier un alpha de Cronbach de satisfaisant (Navarro & Foxcroft, 2020). En effet, si nous observons les coefficients de Cronbach relatifs aux trois grandes compétences, nous pouvons constater qu'ils sont assez faibles et non satisfaisants puisqu'ils sont tous inférieurs à 0.7. Le fait d'avoir modifié l'évaluation initiale de l'EENC de départ peut être une des raisons qui expliquerait ces chiffres.

Seul l'alpha relatif à l'ensemble du test peut être considéré comme satisfaisant et nous permet de qualifier notre test de fiable. Une des pistes que nous avons explorées était de voir si supprimer un item parmi les seize permettrait d'augmenter l'alpha du test. Après analyse, il ressort que sur seize items, seul supprimer l'item 3b aurait permis d'obtenir un alpha très légèrement supérieur. Celui-ci aurait été de 0.71. Néanmoins, nous avons pris la décision de garder cet item, et ce pour deux raisons. Tout d'abord, cet item 3b est lié à l'item 3a. Supprimer l'un aurait donc entraîné la suppression de l'autre. Ensuite, supprimer ces items aurait eu comme conséquence de n'avoir plus que deux items relatifs à la compétence C3 « Transférer », ce qui est trop peu. Enfin, cela n'aurait eu comme effet que d'augmenter l'alpha de Cronbach global du test d'un centième, ce qui est négligeable. Supprimer n'importe quel autre item

n'aurait pas permis d'obtenir un alpha global supérieur. Néanmoins, comme dit précédemment, celui-ci est déjà satisfaisant, ce qui nous permet valider sa consistance interne et de développer nos analyses liées à ce test. Enfin, comme nous l'avons précisé plus tôt (*cf. 3. Question de recherche et hypothèses*), nos hypothèses ne se concentrent pas sur chacune des compétences prises séparément, mais plutôt à l'ensemble du test.

5.2. Résultats du test de mathématique

Dans cette section, nous allons présenter et analyser les données et les résultats relatifs au premier instrument de mesure de notre dispositif, c'est-à-dire le test de mathématique. Pour rappel, il s'agit d'une évaluation sommative à enjeu portant sur les fonctions ayant seize items. Chaque item correspondant à un point, les moyennes présentées sont donc sur seize. Cette évaluation a été administrée de deux manières différentes : l'une mettant en avant des aspects formatifs, de maîtrise (Condition expérimentale / CE) et l'autre mettant en avant des aspects normatifs et de performance (Condition contrôle / CC). Concernant la correction et le codage, chaque bonne réponse récompensait l'élève d'un point tandis que chaque réponse erronée, double réponse ou absence de réponse était codée zéro et ne rapportait pas de point à l'élève.

Les données présentées ci-dessous nous permettront d'apporter des éléments de réponse aux quatre premières hypothèses formulées (Hypothèses 1a, 1b, 1c et 1d). Pour rappel, celles-ci portaient sur les écarts de performance de différentes catégories d'élève au test de mathématique. L'analyse des données et les réponses qu'elles permettent d'apporter à ces premières hypothèses sont exposées ci-dessous à la suite des diverses présentations de résultats.

5.2.1. Une analyse préalable de nos échantillons

Avant d'analyser en détail les résultats des différentes catégories d'élève au test et de comparer les moyennes de celles-ci, nous avons comparé les moyennes de l'ensemble des élèves des deux échantillons en ayant recours au test t de Student pour échantillons indépendants. Ce test nous permet d'évaluer si les moyennes de nos deux groupes (nos deux échantillons) diffèrent significativement ou non entre elles. Nous souhaitons tout de même préciser que dans notre étude, ce t de Student, bien qu'il soit une statistique inférentielle, n'a pas vocation à être inféré à la population, mais plutôt de considérer que la différence observée dans notre échantillon ne semble pas due au hasard.

L'hypothèse nulle (H_0) que nous établissons est celle-ci : les moyennes des deux groupes sont identiques et il n'existe pas de différence significative. Le t-Student nous permet de confirmer ou infirmer cette hypothèse nulle. Notre hypothèse alternative est qu'il y a bien une différence entre les moyennes de nos deux échantillons et que celles-ci ne sont pas égales.

Tableau 7 : Comparaison des deux groupes pour l'ensemble du test de mathématique

		CC (Performance)	CE (Maîtrise)
Test	Moyenne	11.6	12.3
	Écart-type	2.60	2.82
	Variance	6.74	7.94
	Min ; Max	5 ; 16	5 ; 16
	t de Student	-2.85 (<i>p value 0.005</i>)	

Comme l'indique le tableau ci-dessus, l'hypothèse nulle qui suppose une égalité des moyennes entre les deux groupes est à rejeter. En effet, les élèves de la condition CE présentent une moyenne au test supérieure à celle des élèves de la condition CC. Cette différence est jugée significative car la *p value* du *test t de Student* que nous avons effectuée est inférieure à 0.05. Néanmoins, il nous reste à savoir si nous pouvons réellement interpréter ce t de Student et le considérer comme fiable. En effet, pour qu'il soit valide, il est nécessaire de remplir quatre conditions. Tout d'abord, les deux échantillons doivent être indépendants et les variables doivent être à l'échelle métrique. Ces deux premières conditions sont remplies dans notre étude. Ensuite, la variance au sein des groupes doit être similaire. Pour vérifier cette condition, nous avons réalisé un test de Levene. Nous obtenons $F = 3.03$ et une *p value* de 0.082. Notre *p value* étant supérieure à 0.05, nous pouvons conclure que les variances entre les deux groupes des deux conditions sont similaires.

Enfin, pour obtenir un test t de Student valide et précis, il est préférable que les données de chaque échantillon soient normalement distribuées. L'histogramme représentant la densité des données de chaque échantillon ([Annexe 5](#)) montre que la distribution de la CE ne semble pas se distribuer tout à fait normalement. Elle présente une légère asymétrie. Afin de vérifier cette hypothèse de non-normalité, nous avons effectué un test de normalité de Shapiro-Wilk. Nous obtenons $W = 0.967$ et une *p value* de $<.001$. Ayant une *p value* inférieure à 0.05, nous pouvons conclure que notre échantillon ne suit pas une loi normale et que le *test t de Student* calculé préalablement peut donc éventuellement être considéré comme incorrect, car il ne

remplit pas cette condition. Néanmoins, nos deux échantillons étant de taille similaire ($N_{CC} = 245$ et $N_{CE} = 240$) et les variances étant égales, le test t peut résister à une légère asymétrie des données. Toutefois, en cas de non-normalité des données, il est conseillé de réaliser un test U de Mann-Whitney qui tient compte de cette asymétrie. Lorsque nous réalisons ce test U de Mann-Whitney avec les données de nos deux échantillons, nous obtenons une p value de 0.002. À nouveau inférieur au seuil de 0.05, tout comme le t de Student (p value de 0.005), ce résultat nous permet de confirmer qu'il existe bel et bien une différence significative entre les moyennes des élèves ayant suivi la CC et ceux ayant pris part à la CE pour la passation de ce test de mathématique.

5.2.2. Comment qualifier la différence de moyenne entre les échantillons ?

Un dernier élément nous semble important à ne pas négliger : l'ampleur de l'effet. En effet, si les tests de t de Student et U de Mann-Whitney nous ont permis de conclure qu'il existe une différence significative entre les moyennes de nos deux échantillons, ceux-ci ne nous donnent pas d'information sur la force de la différence entre ces moyennes. L'ampleur de l'effet permet quant à elle de qualifier cette différence. Cet indice est calculé à l'aide du d de Cohen. Nous avons réalisé les calculs suivants (comme pour le reste de cette étude) via le site internet Psychometrica.de en utilisant le type d'analyse « Comparison of groups with different effect sizes ».

Tableau 8 : Calcul de l'ampleur de l'effet des deux échantillons

		CC (Performance)	CE (Maîtrise)
Test	Moyenne	11.6	12.3
	Écart-type	2.60	2.82
	N	245	240
	D de Cohen (d)	- 0.26	

Si nous comparons ces résultats avec les seuils habituellement décrétés pour le d de Cohen, nous pouvons conclure que la différence entre les moyennes des deux échantillons est modeste, voire faible. En effet, un d de Cohen se situant autour de 0.2 est considéré comme « faible » ou « small effect » (Navarro & Foxcroft, 2020 ; Lenhard & Lenhard, 2022). Interpréter cette différence entre nos deux échantillons, bien qu'elle n'est peut-être pas nécessairement et uniquement une conséquence de notre dispositif, nous permet de relativiser l'importance pratique du résultat, indépendamment de la taille de l'échantillon. De plus, cela

nous permet d'éviter de tirer des conclusions exagérées sur la seule base de la significativité statistique. Dans la suite de cette partie d'analyse des données, nous allons effectuer un zoom sur les différents groupes qui composent les deux échantillons des deux conditions afin de répondre à nos hypothèses.

5.2.3. Hypothèses 1a, 1b, 1c et 1d

Dans cette section, nous allons tenter de comparer les performances et les écarts de performances des différents groupes qui composent les deux échantillons. Le tableau ci-dessous va nous permettre d'infirmer ou confirmer nos quatre premières hypothèses exposées préalablement. Ces d de Cohen pour échantillons indépendants de N différents ont été calculés grâce au site internet « Psychometrica.de ». Pour rappel, le calcul du d de Cohen étant : moyenne CC moins moyenne CE divisé par la moyenne pondérée des deux écarts-types.

Tableau 9a : Mises en perspective des performances des différents groupes d'élèves

		Test /16	
		CC (performance)	CE (maitrise)
Ensemble Es.	N	245	240
	Moyenne (écart-type)	11.6 (2.60)	12.3 (2.82)
	D de Cohen	-0.26	
Doublants (D)	N	39	42
	Moyenne (écart-type)	10.1 (2.16)	10.5 (2.00)
	D de Cohen	-0.19	
Non-doublants (ND)	N	206	198
	Moyenne (écart-type)	11.9 (2.57)	12.7 (2.82)
	D de Cohen	- 0.3	
Filles (F)	N	108	113
	Moyenne (écart-type)	11.4 (2.58)	11.9 (2.90)
	D de Cohen	-0.18	
Garçons (M)	N	137	127
	Moyenne (écart-type)	11.7 (2.61)	12.6 (2.71)
	D de Cohen	-0.34	

De manière générale, ce tableau nous permet tout d'abord d'établir le constat suivant : dans tous les groupes étudiés et dans les deux variables observées (genre et retard scolaire), la

condition expérimentale (maîtrise) offre des moyennes de performance supérieures à la condition contrôle (performance). Cette observation est à relativiser car les différents d de Cohen calculés sont tous compris entre -0.18 et -0.34, donc à l'avantage de la condition expérimentale axée vers la maîtrise. Selon les repères définis dans la littérature, ces données sont à qualifier de petites, ces différences de moyennes sont faibles, mais perceptibles (Lenhard & Lenhard, 2022).

Hypothèse 1a : Tentons désormais de vérifier notre hypothèse principale. Pour rappel, celle-ci suggérerait que l'écart de performance entre les deux conditions serait davantage marqué pour les élèves en retard que pour les élèves à l'heure. Rappelons également qu'à partir de la littérature scientifique, nous émettions l'hypothèse que la condition expérimentale (CE) serait la plus bénéfique. Cette hypothèse se révèle fausse. En effet, dans notre échantillon, l'écart entre les deux conditions est moins marqué pour les élèves en retard (D) (diff. = +0.4, d de Cohen = -0.19) que pour les élèves à l'heure (ND) (diff. = +0.8, d de Cohen = -0.3). Cette observation suggère que la CE axée vers la maîtrise profite légèrement plus aux non-doublants.

Hypothèse 1c : l'écart de performance entre les deux conditions sera davantage marqué pour les filles que pour les garçons. Cette hypothèse est également à réfuter. Les données récoltées dans le tableau 9a démontrent que l'écart de performance entre les deux conditions n'est pas plus marqué chez les filles. En effet, il est à l'inverse plus présent chez les garçons (diff. de +0.9 contre +0.5 pour les filles) et le d de Cohen est plus important (-0.34 pour les garçons contre -0.18 pour les filles).

Tableau 9b : Mises en perspective des performances des doublants et non-doublants selon la condition du dispositif

		Test /16	
		Doublants (D)	Non-doublants (ND)
CC (performance)	N	39	206
	Moyenne (écart-type)	10.1 (2.16)	11.9 (2.57)
	D de Cohen	-0.72	
CE (maîtrise)	N	42	198
	Moyenne (écart-type)	10.5 (2.00)	12.7 (2.82)
	D de Cohen	-0.82	

Hypothèse 1b : l'écart de performance entre les doublants et les non-doublants serait davantage marqué dans la CC que dans la CE. L'analyse des données obtenues et exposées dans

le tableau ci-dessus nous permet de réfuter cette hypothèse également. En effet, l'écart entre doublants et non-doublants est plus grand dans la CE (écart de 2.2 points et de de Cohen de -0.82) que dans la CC (1.8 points et d de Cohen de -0.72).

Tableau 9c : Mises en perspective des performances des filles et garçons selon la condition

		Test /16	
		Filles (F)	Garçons (M)
CC (performance)	N	108	137
	Moyenne (écart-type)	11.4 (2.58)	11.7 (2.61)
	D de Cohen	-0.17	
CE (maitrise)	N	113	127
	Moyenne (écart-type)	11.9 (2.90)	12.6 (2.71)
	D de Cohen	-0.25	

Hypothèse 1d : l'écart de performance entre les filles et les garçons sera davantage marqué dans la condition contrôle que dans la condition expérimentale. Tout comme les trois hypothèses précédentes, cette hypothèse est également réfutée par les données récoltées. En effet, l'écart de performance est plus grand dans la CE (diff. = +0.7 et d de Cohen de -0.25) que la CC (diff. de +0.3 et d de Cohen de -0.17). Autrement dit, les garçons bénéficient davantage de la CE, axée vers la maitrise que les filles.

5.3. Résultats du questionnaire contextuel

L'objectif de cette partie est de déterminer si la CE, qui constitue une « autre » manière de présenter les évaluations, a un impact sur six variables : le niveau d'anxiété lié aux mathématiques (ANX), le concept de soi en mathématique (SOI) ainsi que les buts d'accomplissement (maitrise-approche = BMA, maitrise-évitement = BME, performance-approche = BPA, performance-évitement = BPE). L'analyse des données de ce questionnaire nous permettra également de répondre aux hypothèses 2, 3 et 4.

Comme précisé préalablement, les six dimensions ont été mesurées à l'aide d'une échelle de Likert allant du score 1 correspondant à « Pas du tout d'accord » au score 4 équivalant à « Tout à fait d'accord ». Chaque réponse d'élève coïncide à une note. Pour connaître la manière dont les étudiants se situent par rapport aux six variables mesurées, nous avons calculé la moyenne des notes obtenues pour chaque construct et pour chaque élève. Un score supérieur à 2.5 indique donc une tendance positive concernant la dimension évaluée. Ci-

dessous sont présentés les résultats, autrement dit, les scores moyens des différentes catégories d'élèves pour chaque variable et pour chaque condition. Le nombre d'individus, l'écart-type ainsi que le d de Cohen pour chaque situation sont également mentionnés dans ce tableau.

Tableau 10 : Mise en perspective des résultats aux variables étudiées au travers du questionnaire contextuel

		SOI		ANX		BMA		BME		BPA		BPE	
		CC	CE	CC	CE	CC	CE	CC	CE	CC	CE	CC	CE
Ens. Es	N	240	239	239	234	238	234	233	233	235	234	238	231
	Moyenne (écart-type)	2.78 (0.58)	2.99 (0.65)	2.49 (0.61)	2.39 (0.68)	2.93 (0.65)	2.91 (0.6)	2.45 (0.68)	2.43 (0.6)	2.24 (0.68)	2.36 (0.75)	2.86 (0.54)	2.76 (0.62)
	D de Cohen	0.34		-0.16		-0.03		-0.03		0.17		-0.17	
D	N	37	41	38	41	38	39	38	39	37	39	38	37
	Moyenne (écart-type)	2.62 (0.55)	2.9 (0.64)	2.54 (0.57)	2.42 (0.63)	2.81 (0.7)	2.85 (0.6)	2.32 (0.73)	2.39 (0.46)	2.03 (0.7)	2.39 (0.73)	2.74 (0.59)	2.68 (0.48)
	D de Cohen	0.45		-0.2		0.06		0.12		0.5		-0.11	
ND	N	203	198	201	193	200	195	195	194	198	195	200	194
	Moyenne (écart-type)	2.81 (0.58)	3 (0.65)	2.48 (0.62)	2.38 (0.69)	2.95 (0.63)	2.92 (0.6)	2.48 (0.66)	2.43 (0.62)	2.28 (0.68)	2.35 (0.76)	2.88 (0.53)	2.78 (0.64)
	D de Cohen	0.31		-0.15		-0.05		-0.08		0.1		-0.17	
F	N	107	112	105	111	104	110	102	111	102	111	104	109
	Moyenne (écart-type)	2.71 (0.56)	2.93 (0.73)	2.64 (0.63)	2.56 (0.67)	2.93 (0.68)	2.98 (0.61)	2.52 (0.73)	2.47 (0.58)	2.25 (0.68)	2.41 (0.8)	2.94 (0.53)	2.91 (0.57)
	D de Cohen	0.34		-0.12		0.08		-0.08		0.22		-0.06	
M	N	133	127	134	123	134	124	131	122	133	123	134	122
	Moyenne (écart-type)	2.84 (0.59)	3.03 (0.57)	2.37 (0.56)	2.23 (0.66)	2.93 (0.62)	2.84 (0.58)	2.39 (0.63)	2.39 (0.62)	2.23 (0.69)	2.32 (0.72)	2.80 (0.54)	2.64 (0.62)
	D de Cohen	0.33		-0.23		-0.15		0		0.13		-0.28	

Tout d'abord, une observation globale nous permet de dire que l'ensemble des ampleurs de l'effet sont assez faibles, voire parfois nulles. L'ampleur de l'effet la plus importante étant celle entre les doublants pour la variable BPA qui obtient un d de Cohen de 0.5. Pour interpréter ces d de Cohen, nous nous référons aux seuils prônés par les chercheurs où selon eux, le seuil de 0.2 évoque une ampleur de l'effet « faible » et le seuil de 0.5 correspond à une ampleur de l'effet qualifié de « moyenne » (Navarro & Foxcroft, 2020 ; Lenhard & Lenhard, 2022). Ces faibles tailles de l'effet nous laissent penser que les différences et les écarts de réponse entre la CC et la CE sont relativement faibles sauf concernant la variable SOI qui de manière générale semble relever des ampleurs de l'effet plus importantes. Même si ces données semblent d'ampleur plutôt faible, nous analyserons celles-ci et/ou les observerons de manière plus indicative lorsque cela sera plus approprié.

5.3.1. Hypothèses 2a, 2b et 2c

Pour rappel, les hypothèses 2a, 2b et 2c traitent la dimension de l'anxiété liée aux mathématiques. Nous faisons l'hypothèse, à partir des éléments dégagés dans la revue de la littérature, que les filles et les doublants s'imagineraient, en CC axée sur la performance, moins compétents que les autres, ce qui provoquerait l'apparition de pensées intrusives, de stress et d'anxiété (Goudeau, 2023 ; Lafontaine & Toczek, 2023 ; Muller & Butera, 2007).

Notre tableau descriptif nous apprend en premier lieu que de manière générale, les élèves éprouvent un niveau d'anxiété relativement neutre. En effet, le tableau nous permet d'observer que les scores concernant cette variable ANX sont toujours proches du point médian qui est de 2.5. Seuls les garçons en CE semblent se distinguer un peu plus de ce point médian avec un score de 2.23. En outre, quel que soit le groupe, les élèves s'évaluent plus anxieux vis-à-vis des mathématiques dans la CC que dans la CE. Concernant les deux publics cibles que visent nos hypothèses (filles et doublants), nous pouvons observer que ceux-ci également se disent plus anxieux en ce qui concerne les mathématiques dans la CC que dans la CE. En effet, le score moyen des filles passe de 2.64 en CC à 2.56 en CE et les doublants passent d'un score moyen de 2.54 en CC à 2.42 en CE. Cependant, nous nous devons de rester prudents et mesurés dans les conclusions que nous pouvons tirer de ces résultats, car comme l'indiquent les différentes ampleurs de l'effet, la différence est faible (de de Cohen de -0.12 pour les filles et de -0.2 pour les doublants). Néanmoins, nous pouvons confirmer avec prudence (étant donné les taux des ampleurs de l'effet) nos hypothèses 2a, 2b et 2c selon lesquelles l'ensemble des élèves, les filles, mais aussi les doublants sont davantage anxieux vis-à-vis des mathématiques en CC que lorsqu'ils sont placés dans la CE.

5.3.2. Hypothèses 3a, 3b et 3c

Les hypothèses 3a, 3b et 3c interrogent la dimension du concept de soi en mathématique et se concentrent sur les filles et les doublants. En effet, nous émettions l'hypothèse que l'ensemble des élèves, mais aussi les filles et les élèves en retard enregistreraient des scores moyens plus faibles en CC qu'en CE concernant le SOI. Avant de confirmer ou infirmer ces hypothèses, nous pouvons avant tout souligner que ce construct SOI est celui qui obtient des ampleurs de l'effet les plus importantes parmi l'ensemble des dimensions étudiées dans notre dispositif. En effet, les ampleurs de l'effet des d de Cohen varient entre 0.31 et 0.45. Bien que ces différences ne soient pas à considérer comme fortes, il reste intéressant de noter que c'est cette dimension SOI qui enregistre celles les plus élevées.

Concernant les groupes visés par nos hypothèses (notamment les filles et les doublants), nous pouvons déclarer à partir des données obtenues, que ceux-ci obtiennent des scores moyens plus élevés dans la CE que dans la CC. Autrement dit, l'ensemble des élèves, mais aussi les filles et les doublants estiment avoir un concept de soi plus positif lorsque l'évaluation est orientée vers la maîtrise plutôt que vers la performance. En effet, les filles enregistrent un score moyen de 2.71 en CC et 2.93 en CE. Les doublants quant à eux, passent d'un score moyen de 2.62 en CC à 2.9 en CE. À titre indicatif, nous pouvons également constater que les garçons et les non-doublants expriment également des concepts de soi liés aux mathématiques plus élevés en CE qu'en CC. En conclusion, l'analyse des données relatives au construct SOI nous permet de confirmer nos trois hypothèses relatives à ce même construct.

5.3.3. Hypothèses 4a, 4b et 4c

Enfin, les hypothèses 4a, 4b et 4c interrogent les buts d'accomplissement. Concernant ces hypothèses et au vu des ampleurs de l'effet de ces quatre variables liées aux buts d'accomplissement (très faibles dans l'ensemble, voire parfois nulles), nous souhaitons davantage analyser ces données de manière indicative afin d'observer éventuellement des tendances qui se dégageraient, mais ces résultats ne nous permettent pas de valider clairement ces hypothèses. Néanmoins, nous faisons l'hypothèse (4a) qu'en CC, les élèves seraient davantage enclins à poursuivre des buts de performance. Sur base des données récoltées, il est très difficile de valider ou réfuter cette hypothèse, car comme le tableau descriptif le montre ci-dessous, en ce qui concerne les buts d'accomplissement, les ampleurs de l'effet entre les deux conditions sont très faibles, voire nulles. Nous préférons par conséquent ne pas donner de réponse trop ferme concernant cette hypothèse. Nous pouvons plutôt déclarer que dans notre échantillon, les élèves poursuivent de manière semblable des buts de performance dans les deux conditions.

Tableau 11 : Mises en perspective des résultats relatifs aux buts d'accomplissement

		CC (performance)	CE (maîtrise)
Buts de maîtrise (BM)	N	233	232
	Moyenne (écart-type)	2.69 (0.57)	2.67 (0.47)
	D de Cohen	-0.04	
Buts de performance (BP)	N	235	230
	Moyenne (écart-type)	2.55 (0.45)	2.56 (0.5)
	D de Cohen	0.02	

Les hypothèses 4b et 4c s'intéressent quant à elles au groupe des doublants. Selon ces hypothèses, ceux-ci devraient suivre davantage de buts de performance-évitement, mais moins de buts de maîtrise-approche en CC. À partir des données exposées dans le tableau descriptif numéro 10 (cf. *Tableau 10 : Mise en perspective des résultats aux variables étudiées au travers du questionnaire contextuel*), nous pouvons affirmer qu'en effet les doublants de notre échantillon placés en CC poursuivent davantage de BPE (2.74 en CC et 2.68 en CE, d de Cohen de -0.11). En ce qui concerne la seconde hypothèse, les doublants de notre échantillon obtiennent un score moyen BMA de 2.80 en CC et de 2.85 en CE (d de Cohen de 0.06). Néanmoins, nous nous devons à nouveau d'être prudents quant à ces conclusions, car les ampleurs de l'effet de ces variables restent relativement faibles, voire négligeables. Enfin, nous nous devons également de rappeler que l'alpha de Cronbach relatif à la variable but de performance-évitement n'était pas suffisamment bon, ce qui entraîne davantage à la prudence et nous amène une nouvelle fois à considérer ces résultats de manière indicative.

5.4. Interprétation et discussion des résultats

En préambule, il faut reconnaître que réfuter quatre hypothèses formulées dans le cadre de notre étude peut paraître quelque peu intrigant et surprenant (il s'agit des hypothèses 1a, 1b, 1c et 1d). Dès lors, nous nous posons la question de savoir si nos hypothèses ont été mal formulées et mal pensées en amont ou si les résultats de nos échantillons vont à l'encontre de ce que suppose la littérature scientifique. Nous tentons ci-dessous de comprendre quelles interprétations et quelles conclusions nous pouvons retirer de ces résultats en nous replongeant également dans ce que dit la littérature scientifique.

En premier lieu, avant de discuter de la non-confirmation des hypothèses relatives aux performances obtenues à l'évaluation, nous souhaitons reprendre un angle de vue général sur les résultats obtenus au test de mathématique par les élèves. Comme nous l'avons dit précédemment, nous observons que tous les groupes d'élèves, variables du genre et du retard scolaire comprises, obtiennent de meilleurs résultats dans la CE qui insiste sur les aspects formatifs de l'évaluation en veillant à mettre en avant des objectifs d'apprentissage et de maîtrise. En effet, comme nous l'avons expliqué dans la formulation de nos hypothèses dans la partie méthodologie, nous pensions, en accord avec la littérature scientifique, que l'évaluation sommative normative allait engendrer toute une série de menaces et conséquences néfastes pour la majorité des élèves. Nous pouvons citer comme cause de baisse de performance la poursuite de buts de performance évitement, le rappel d'épisodes douloureux et l'apparition

d'émotions négatives, les stéréotypes négatifs visant les filles et les doublants ou encore la comparaison sociale entre élèves. Tous ces potentiels dangers et possibles causes de baisse de performance peuvent avoir entraîné une baisse des résultats de l'ensemble des catégories d'élève dans la CC comme nous l'imaginions et comme la littérature le suggérait. Là où nos quatre hypothèses semblent être ébranlées, c'est en réalité davantage dans les différences d'écart de performance de certaines catégories d'élèves entre les deux conditions qui posent question. En effet, nous émettions comme hypothèse que les filles et les doublants obtiendraient une différence de performance plus importante entre les deux conditions comparativement aux garçons et aux non-doublants. En réalité, c'est le contraire qui apparaît dans notre étude, c'est la CE axée vers la maîtrise qui bénéficie davantage aux non-doublants. Il faut néanmoins préciser que cette différence d'écart de performance entre ces catégories d'élève est très faible. Ces données nous laissent penser qu'il n'y a pas d'effet clair marqué par les deux conditions.

La question que nous nous posons est pourquoi les garçons et les non-doublants obtiennent un écart de performance entre les deux conditions plus important que les filles et les doublants. Il est difficile de répondre à cette question, mais une des pistes que nous pouvons avancer concerne l'effet que la forme de ce dispositif peut entraîner. En effet, pour rappel, cette évaluation, dans les deux conditions, reste impressionnante et peut perturber. Malgré sa programmation annoncée d'emblée, elle intervient parfois de manière surprenante, car certaines classes ont vu la matière parfois plusieurs mois avant la passation. Ensuite, la mise en place de ce test peut faire peur et raviver des souvenirs douloureux pour certains. Sa mise en page est celle des épreuves non certificatives, mais elle est également très semblable à celle des épreuves certificatives. De plus, l'enseignant suit un guide de passation, il doit lire des consignes très précises. Les élèves doivent aussi compléter un court questionnaire contextuel en amont du test en répondant à des questions personnelles. Tous ces facteurs peuvent engendrer du stress, mettre en évidence des stéréotypes négatifs, engendrer des pensées intrusives, rappeler des épisodes douloureux et in fine de possibles baisses de performance (Lafontaine & Toczek, 2023 ; Fresson & Dardenne, 2019 ; Goudeau, 2023). Des groupes plus fragiles comme les doublants et les filles peuvent donc être davantage mis sous pression, perturbés que les garçons et les non-doublants et donc moins bien performer (Toczek & Souchal, 2017). Ceci peut avoir eu lieu dans les deux conditions. Une des possibilités est aussi que ces éléments aient été plus prégnants que la consigne qui, elle, différait d'une condition à l'autre. Pour rappel, c'est cette consigne (inspirée de celle proposée par Souchal et ses collaborateurs) et uniquement cette consigne qui différenciait les deux conditions de notre étude. Cependant, il est utile de noter que Souchal et

ses collègues (2014) en utilisant cette consigne également obtiennent des différences plus importantes entre les filles et les garçons. Même si le dispositif global entre notre présente étude et celle réalisée en 2014 par ses chercheurs se ressemble, plusieurs différences peuvent être pointées. En effet, dans l'étude de Souchal et ses collaborateurs (2014), les élèves étaient une année plus âgés, l'évaluation portait sur les sciences, l'heure de rappel précédait obligatoirement le test, les classes étaient divisées en deux groupes correspondant à deux des trois conditions et placées dans deux salles distinctes et enfin, les expérimentatrices elles-mêmes dirigeaient le dispositif. Voici une série de facteurs absents de notre étude qui peuvent nous aider à mieux envisager et considérer cette différence de résultats entre les deux recherches.

En outre, ces résultats issus du test peuvent être interprétés à la lumière des données récoltées au travers du questionnaire contextuel relatif aux variables conception de soi, anxiété et buts d'accomplissement. Les interprétations faites concernant ces trois variables peuvent apporter un éclairage nouveau sur ces absences de différences fortes entre les deux conditions et le rejet de nos quatre hypothèses (1a, 1b, 1c et 1d). En effet, l'analyse des résultats du questionnaire contextuel met par exemple en avant que les élèves, toutes catégories confondues, ne se disent que légèrement plus anxieux en condition contrôle. Ceux-ci exposent également une conception de soi davantage positive en condition axée sur la maîtrise, mais la différence entre les conditions reste à nouveau faible. Ces faibles différences entre les deux conditions observées concernant ces deux variables peuvent en partie expliquer les faibles différences constatées en ce qui concerne l'évaluation de mathématique. Des différences importantes entre les deux conditions concernant les variables SOI, ANX ou BUTS auraient peut-être engendré également des différences plus importantes entre les deux conditions en ce qui concerne le test. Finalement, les résultats de notre étude concernant le questionnaire indiquent dans l'ensemble des effets soit neutres soit légèrement positifs en condition expérimentale. Cette conclusion est semblable à celle réalisée pour l'évaluation de mathématique.

En conclusion, la CE bénéficie légèrement aux non-doublants et aux garçons mais les filles et les doublants obtiennent également de meilleurs résultats dans la CE par rapport à la CC. Ce constat peut être expliqué par les raisons évoquées ci-dessus. Il semble que la différence de consignes entre les deux conditions ait provoqué moins de différences que ce que nous imaginions dans nos hypothèses car comme démontré précédemment, les différences observées concernant l'anxiété, le concept de soi ou les buts d'accomplissements entre les conditions sont soit faibles soit nulles. Enfin, nous pouvons conclure, sur base des données récoltées dans notre

étude, que bien que cette manière de présenter cette évaluation en l'axant vers des éléments de maîtrise n'ait pas engendré des différences de performances importantes, elle n'a pas entraîné de baisses de résultats et n'a nui à aucune catégorie d'élèves de notre échantillon (dont les élèves en retard scolaire qui sont le public cible de cette étude).

6. Limites et perspectives

Avant de mettre en avant les diverses projections qui pourraient être réalisées à la suite de cette recherche, nous nous devons de reconnaître les limites et les biais potentiels de celle-ci. Bien qu'un maximum de dispositions aient été prises afin de limiter ces biais, certaines limites nous semblent importantes à préciser et prendre en compte. La première limite que nous souhaitons évoquer concerne nos échantillons et l'équivalence entre ceux-ci. Afin de maximiser les chances d'équivalence entre nos deux groupes, nous avons réparti aléatoirement les classes et veillé à ce que les classes rassemblant des options particulières soient distribuées équitablement dans les deux conditions. De plus, tous les élèves ont le même nombre d'heures de mathématiques par semaine, dans la majorité des cas les enseignants ont une classe dans chaque condition et le nombre de classes et d'individus est presque équivalent dans les deux cas. Bien qu'ils soient suffisamment larges, la réalité de terrain nous a obligés à modifier quelques-uns des critères (notamment sur le type d'enseignement et sur le nombre de classes par professeur). Nos échantillons restent également des échantillons de convenance, ce qui nous amène également à la prudence et à ne pas généraliser nos conclusions. C'est également parce que nous ne pouvons assurer une équivalence parfaite au départ entre nos deux échantillons qu'il nous revient d'être prudents dans les comparaisons brutes effectuées entre la condition contrôle et la condition expérimentale. C'est également pour cette raison que nous avons formulé des hypothèses en termes d'écart de performances entre les conditions.

Bien qu'un ensemble de mesures aient été prises (entretiens, questions-réponses, guide de passation...), le chercheur n'étant pas dans les classes, la façon dont les dispositifs ont été mis en pratique par les enseignants peut également être un potentiel biais. Enfin, nous ne pouvons négliger les éventuels biais de réponses apparus dans le cadre du questionnaire contextuel. Ceux-ci sont définis par Paulhus (1991) comme étant « une tendance systématique à répondre à un ensemble d'items d'un questionnaire sur base autre que le contenu spécifique de l'item » (p.17, cité par Lafontaine, 2023). Enfin, il serait pertinent de prendre en compte un potentiel effet de fatigue. Effectivement dans notre étude, les élèves passent une période entière

à travailler individuellement sans interaction. Nous pouvons donc imaginer que le moment venu de répondre au questionnaire contextuel, les élèves soient impatients et répondent à la « va vite » aux items afin d'en être débarrassés. Voici donc une série de biais ou de limites que nous pouvons déceler dans notre étude.

Outre ces limites, nous souhaitons dresser une série de perspectives que nous entrevoyons concernant cette étude qui permettraient d'aller un pas plus loin dans les conclusions de celle-ci. Dans un premier temps, il serait pertinent de modifier certains critères d'échantillonnage et de passation du dispositif afin d'obtenir une équivalence d'échantillons et des conditions de passation encore plus satisfaisantes (ex : s'intéresser au niveau socio-économique des écoles sélectionnées, ne choisir que des classes d'un même type d'enseignement, etc.). Cela dit, nous pensons qu'il serait intéressant d'approfondir l'analyse des données récoltées. Il serait par exemple pertinent d'observer les liens potentiels entre les scores des élèves aux différents constructs du questionnaire contextuel et leurs performances au test. Il serait également possible de combiner plusieurs variables (par exemple le retard et le genre) dans l'analyse des résultats au test. Autrement dit, en plus d'analyser les performances des doublants et des filles séparément, il serait par exemple intéressant de comparer celles des filles doublantes afin de voir si ces variables combinées rendent ces élèves encore plus sensibles aux différents contextes évaluatifs ou non. Une des raisons qui ne nous a pas permis de réaliser ces recherches complémentaires est le nombre insuffisant de sujets pour ces catégories (nous n'avons par exemple que dix filles doublantes dans notre condition contrôle).

Enfin, les conclusions de notre recherche, comme les différentes études menées sur l'influence du contexte évaluatif pourraient amener les écoles et les enseignants à repenser la manière dont ils conçoivent et mettent en place leurs évaluations. En effet, présenter les évaluations en les orientant vers des objectifs de maîtrise de manière durable et continue dans une école aurait sans doute des résultats encore plus probants et positifs.

7. Conclusion

Face à une situation d'évaluation, l'élève n'est pas seul dans sa bulle, dénué d'émotions et impassible à ce qui l'entoure et/ou à son vécu. Diverses études ont révélé que le contexte évaluatif dans lequel se tient cette activité est crucial (Lafontaine & Toczek, 2023). En réalité, toute une série de phénomènes psychosociaux entrave les performances comme le type d'évaluation, la motivation, les interactions sociales, les expériences passées ou encore la

menace du stéréotype (Monteil & Huguet, 2011). C'est à partir de ce point de départ que nous avons tenté d'explorer cette problématique de l'influence du contexte évaluatif sur les performances mathématiques selon le retard scolaire. Ce que nous estimons important de retenir de notre étude, est le fait que nous avons présenté notre évaluation sommative à enjeu en l'orientant vers des aspects formatifs et de maîtrise, cela n'a en aucun cas nui aux performances des élèves, mais au contraire les a légèrement améliorées. En effet, toutes les catégories d'élèves (doublants/non-doublants, filles/garçons) obtiennent des performances supérieures en condition expérimentale lorsque l'évaluation est tournée vers des objectifs de maîtrise. De plus, en situation expérimentale, les élèves ont à nouveau exprimé des ressentis soit relativement semblables qu'en situation contrôle (lorsque les ampleurs de l'effet étaient faibles voire nulles) soit légèrement plus positifs concernant les variables étudiées. Les élèves de notre échantillon éprouvent une meilleure conception de soi et se sentent moins anxieux lorsque l'évaluation est présentée comme ayant des bénéfices en termes d'apprentissage.

La question sous-jacente de notre étude est de savoir s'il est possible de concevoir et mettre en place des évaluations qui sont moins préjudiciables à certaines catégories d'élèves (Butera, 2023). C'est-à-dire des évaluations qui réduisent les inégalités, bénéficient aux élèves les plus fragiles et stéréotypés (dont font partie les élèves en retard) tout en ne nuisant pas à d'autres catégories d'étudiants. Cette idée peut paraître utopique, mais il ne faut pas sous-estimer le rôle central que joue le processus évaluatif dans la vie étudiante (Butera, 2023). De ce fait, il nous semble juste de tout de même s'y intéresser comme d'autres chercheurs l'ont fait précédemment (Toczek et Souchal par exemple). De par la spécificité de notre public ciblé que sont les doublants (le nombre important dans notre enseignement, les stéréotypes auxquels ils doivent faire face, etc.), nous espérons apporter un éclairage complémentaire dans ce domaine si important de l'école que sont l'évaluation et les stéréotypes liés au redoublement. Enfin, comme nous l'avons esquissé précédemment, nous sommes convaincus que ce domaine requiert encore des investigations plus approfondies. Cette étude peut servir de point de départ à des recherches supplémentaires, à plus large échelle et suivant des critères encore davantage précis.

Bibliographie

Appel, M. & Kronberger, N., (2012). Stereotypes and the achievement gap: stereotype threat prior to test taking. *Educational Psychology Review*, 24(4), 609-635. <https://dx.doi.org/10.1007/s10648-012-9200-4>

Autin, F., Batruch, A. & Butera, F. (2019). The Function of Selection of Assessment Leads Evaluators to Artificially Create the Social Class Achievement Gap. *Journal of Educational Psychology*, 111 (4), 717-735. [doi: 10.1037/edu0000307](https://doi.org/10.1037/edu0000307).

Bandura, A. (1982). The self and mechanisms of agency; In J. Suls (dir.), *Psychological perspectives on the self* (1), 3-40. Hillsdale, NJ: Erlbaum.

Baye, A. (2021). *Introduction aux sciences de l'éducation et de la formation* [document inédit]. Université de Liège.

Baye, A., Bricteux, S., Crépin, F., Demonty, I., Hindryckx, G., Matoul, A., Quittre, V. (2023). Résultats de PISA 2022 en Fédération Wallonie-Bruxelles. *Les Cahiers des Sciences de l'Éducation*, 46.

Blouin, Y., (1985). *La réussite en mathématiques au collégial : le talent n'explique pas tout*. Rapport de recherche publié au Cégep François-Xavier-Garneau.

Butera, F. (2023) Quels sont les effets psychosociaux de l'évaluation normative ? dans *Conférence de consensus du Cnesco l'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 5-12). Cnesco-Cnam.

Crahay, M. (2019). *Peut-on lutter contre l'échec scolaire ?* (4^e éd.). Bruxelles : De Boeck Supérieur.

Da Fonseca, D., Cury, F., Bailly, D. & Rufo, M. (2004). Théories implicites de l'intelligence et buts d'accomplissement scolaire. *Annales Medico Psychologiques*, 162, 703-710.

Dardenne, B. (2022). *Psychologie sociale* [document inédit]. Université de Liège.

Darnon, C., Buchs, C. & Butera, F. (2006). Buts de performance et de maîtrise et interactions sociales entre étudiants : la situation particulière du désaccord avec autrui. *Revue française de pédagogie*, 2 (155), 35-44. [doi: 10.4000/rfp.84](https://doi.org/10.4000/rfp.84)

Darnon, C. & Butera, F. (2005). Buts d'accomplissement, stratégies d'étude, et motivation intrinsèque : présentation d'un domaine de recherche et validation française de l'échelle d'Elliot et McGregor (2001). *L'année psychologique*, 105 (1), 105-131.

Dierendonck, Ch. & Cavaco, C. (2017). Évaluer : pour contrôler ou pour apprendre ? *Evaluer. Journal international de Recherche en Education et Formation*, 3(1-2), 99-100.

- Dupont, V., Demonty, I., Quittre, V. & Fagnant, A. (2024). TIMSS 2023. *Note de synthèse*.
- Dutrévis, M., Crahay, M. (2013). Redoublement et stigmatisation : conséquences pour l'image de soi des élèves. *L'Orientation scolaire et professionnelle*, 42. [Doi: 10.4000/osp.4150](https://doi.org/10.4000/osp.4150)
- Dutrévis, M., (2015). Menace du stéréotype et groupes ethno-raciaux minoritaires : quel poids sur les performances des élèves ?. *Revue française de pédagogie*, 191, 61-72.
<https://doi.org/10.4000/rfp.4751>
- Fagnant, A. (2023). *Les pratiques d'évaluation en classe : des compétences professionnelles pour soutenir l'apprentissage des élèves*. Cnesco-Cnam.
- Fédération Wallonie-Bruxelles. (2022). Les Indicateurs de l'enseignement. Dans *Enseignement.be*. Fédération Wallonie-Bruxelles - Administration générale de l'Enseignement.
- Florin, A., Tricot, A., Chesné J.-F., Piedfer-Quêne, L., Simonin-Kunerth, M., (2023). *Dossier de synthèse : L'évaluation en classe, au service de l'apprentissage des élèves*. Cnesco-Cnam.
- Fresson, M. & Dardenne, B. (2019). Les redoublants, victimes de stéréotypes dévalorisants ? dans Crahay, M. (dirs.), *Peut-on lutter contre l'échec scolaire ?* (pp. 231-260). De boeck Supérieur.
- Galand, B. & Grégoire, J. (2000). L'impact des pratiques scolaires d'évaluation sur les motivations et le concept de soi. *L'orientation scolaire et professionnelle*, 29 (3), 431-452.
- Goudeau, S. (2023). En classe, quels effets peuvent avoir les interactions à dimension évaluative sur les élèves en termes d'inégalités ? dans *Conférence de consensus du Cnesco l'évaluation en classe, au service de l'apprentissage des élèves: Notes des experts* (pp. 93-100). Cnesco-Cnam.
- Huguet, P. (2018). Apprendre en groupe : la classe dans sa réalité sociale et émotionnelle. Dans Bourgeois, E. & Chapelle, G. (dirs.), *Apprendre et faire apprendre* (pp. 169-183). PUF
- Jaegers, D. & Lafontaine, D. (2018). Perceptions par les élèves du climat de soutien en mathématiques : validation d'échelles et étude des différences selon le genre en 5e secondaire. *Mesure et évaluation en éducation*, 41 (2), 97–130. <https://doi.org/10.7202/1059174ar>
- Lafontaine, D., Bricteux, S., Hindryckx, G., Matoul, A., & Quittre, V. (2019). Performances des jeunes de 15 ans en lecture, mathématiques et sciences. *Premiers résultats de PISA 2018 en Fédération Wallonie-Bruxelles*.
- Lafontaine, D. & Toczec-Capelle, M.-C. (2023). *L'évaluation en classe au service de l'apprentissage des élèves : rapport de synthèse*. Cnesco-Cnam.
- Lafontaine, D. (2018). *Construction de questionnaires* [document inédit]. Université de Liège.

Lenhard, W. & Lenhard, A. (2022). *Computation of effect sizes*. Consulté sur https://www.psychometrica.de/effect_size.html. (Consulté le 13 juin 2025)

Monteil, J.M. & Huguet, P. (2001). The Social Regulation of Classroom Performances: a Theoretical Outline. *Social Psychology of Education*, 4, 359-372. <https://doi.org/10.1023/A:1011345022870>

Muller, D. & Butera, F. (2007). The Focusing Effect of Self-Evaluation Threat in Coaction and Social Comparison. *Journal of Personality and Social Psychology*, 93 (2), 194-211. doi: [10.1037/0022-3514.93.2.194](https://doi.org/10.1037/0022-3514.93.2.194).

Navarro, D. & Foxcroft, D. (2020). *Apprentissage des statistiques avec Jamovi : un tutorial pour les étudiants en psychologie et autres débutants* (version 0.70.2). <https://hal.science/hal-02335912v2>

OECD. (2016). *PISA 2015 Results. Excellence and equity in education* (Vol. 1). Paris: OECD Publishing.

Pajares, F. & Miller, D.M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86, 193-203.

Paquay, L. (2006). Introduction au-delà des cloisonnements entre divers types de recherche, quels critères de qualité ? In, L. Paquay, M. Crahay & J-M. De Ketele (eds.), *l'analyse qualitative en éducation, des pratiques de recherche aux critères de qualités*, 13-29. Bruxelles, Belgique: De Boeck.

Pasquini, R. (2023). À quelles conditions développer des pratiques de notation au service des décisions des enseignants et des apprentissages des élèves ? dans *Conférence de consensus du Cnesco l'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 136-147). Cnesco-Cnam.

Sénéchal, E. (1997). *La relation entre le concept de soi et le rendement scolaire chez des élèves de troisième année du primaire*. Mémoire en sciences de l'éducation non publié, Université Laval, Québec.

Schwab, S. & Held, L. (2020). Different worlds: confirmatory versus exploratory research. *Significance*, 08-09. <https://doi.org/10.1111/1740-9713.01369>

Slavin, R. E. (2007). *Education Research in an Age of Accountability*. Boston, MA: Pearson Education, 20 – 21.

Souchal, C., Toczek, M.C., Darnon, C. Smeding, A., Butera, F., & Martinot, D. (2014). Assessing Does Not Mean Threatening: Assessment as a key determinant of girls' and boys' performance in a science class. *British Journal of Educational Psychology*, 84(1), 125-136. <https://doi.org/10.1111/bjep.12012>

Souchal, C. & Toczec, M-Ch. (2010). Buts de réussite, conceptions de l'intelligence, différences de performances liées à l'appartenance socio-économique des élèves : de nouvelles hypothèses explicatives ? *Les Sciences de l'Éducation – Pour l'Ère Nouvelle*, 43, 13-35. Doi : [10.3917/lsdle.431.0013](https://doi.org/10.3917/lsdle.431.0013)

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 69 (5), 797-811. <https://doi.org/10.1037/0022-3514.69.5.797>

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613–629.

Toczec, M-Ch. & Souchal, C. (2017). Le pouvoir des contextes évaluatifs. *Evaluer. Journal international de Recherche en Education et Formation*, 3 (1-2), 21-35.

Vilette, B. (2017). L'anxiété mathématique apparaît-elle au début des apprentissages scolaires ? *Enfance*, 4(4), 513-519. <https://doi.org/10.3917/enf1.174.0513>

Weissgerber, S.C., Grünberg, C., Neufeld, L. et al. (2022). The interplay of math anxiety and math competence for later performance. *Social Psychology in Education*, 25, 977–1002. <https://doi.org/10.1007/s11218-022-09700-y>

Table des tableaux

Tableau 1 : Répartition des élèves de l'échantillon dans les deux conditions du dispositif en fonction du genre et du retard scolaire	37
Tableau 2 : Grille de description et de comparaison des échantillons des deux conditions.....	38
Tableau 3 : Présentation simplifiée des étapes concrètes du dispositif de cette recherche.....	39
Tableau 4 : Consignes données par les enseignants selon la condition du dispositif.....	41
Tableau 5 : Alphas de Cronbach des constructs étudiés dans le questionnaire contextuel.....	46
Tableau 6 : Alphas de Cronbach relatifs au test de mathématique.....	47
Tableau 7 : Comparaison des deux groupes pour l'ensemble du test de mathématique.....	49
Tableau 8 : Calcul de l'ampleur de l'effet des deux échantillons.....	50
Tableau 9a : Mises en perspective des performances des différents groupes d'élèves au test de mathématique.....	51
Tableau 9b : Mises en perspective des performances au test des doublants et non-doublants selon la condition du dispositif.....	52
Tableau 9c : Mises en perspective des performances au test des filles et des garçons selon la condition du dispositif.....	53
Tableau 10 : Mises en perspective des résultats des catégories d'élèves aux variables étudiées au travers du questionnaire contextuel.....	54
Tableau 11 : Mises en perspective des résultats relatifs aux buts d'accomplissement.....	56

Table des annexes

ANNEXE 1 : Guide de passation – Manuel de l’enseignant.....	71
ANNEXE 2 : Points de matière à aborder pour la période de rappel.....	83
ANNEXE 3 : Test de mathématique.....	85
ANNEXE 4 : Questionnaire contextuel.....	96
ANNEXE 5 : Histogramme de densité des deux échantillons.....	101
ANNEXE 6 : Lettre aux enseignants.....	102



DOSSIER DE L'ENSEIGNANT
Projet de Mémoire – 2024

Introduction

Nous souhaitons tout d'abord vous remercier d'avoir accepté d'être participant pour cette recherche de mémoire. Ce projet de recherche s'intéresse à une dimension parfois sous-estimée à l'école, mais bien présente dans notre enseignement : l'évaluation. Il semble exister dans notre société une croyance commune très ancrée qui consiste à justifier les échecs des élèves aux évaluations uniquement par leurs capacités intrinsèques (Toczek & Souchal, 2017). Toczek et Souchal (2017), ainsi que d'autres auteurs avant elles, s'efforcent de démontrer qu'une série de facteurs issus du contexte évaluatif peuvent être une source de menace pour certains élèves et donc également jouer un rôle dans les résultats des élèves.

L'étude du mécanisme de la menace du stéréotype montre par exemple que des signaux du contexte évaluatif peuvent dans certains cas faire réussir les élèves et dans d'autres les faire échouer selon que ceux-ci subissent ce phénomène ou non (Lafontaine & Toczek, 2023). Il nous semble crucial de prendre conscience des contextes évaluatifs qui augmentent ou diminuent cette menace du stéréotype et qui ont des répercussions conséquentes sur les performances des élèves.

Ces réflexions nous ont donc amenés à nous poser la question suivante : « dans quelle mesure l'objectif d'une évaluation sommative avec enjeu peut-il influencer les performances des élèves en retard selon qu'elle soit axée sur la performance (c'est-à-dire une évaluation ayant pour but de principalement comparer et sélectionner les élèves) ou qu'elle soit axée sur la maîtrise (c'est-à-dire une évaluation présentée comme ayant également un but d'apprentissage) ? ».

1 Déroulement du dispositif

Concrètement, les élèves passeront un test de mathématiques qui sera noté et que vous pourrez utiliser comme une évaluation traditionnelle qui compte pour la note finale du bulletin. Les items de ce test portent sur les chapitres UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré). L'idée est de manipuler le contexte évaluatif de deux manières différentes afin de voir si certaines manières de présenter l'évaluation peuvent être davantage bénéfiques ou préjudiciables pour les élèves en retard scolaire. Nous comparerons la performance des élèves en situation évaluative axée sur la performance (situation 1) à celle en situation d'évaluation axée sur la maîtrise (situation 2). Les deux groupes passeront le même test.

L'expérimentation nécessitera deux heures de cours pour chaque classe. La première heure consiste en un rappel, un cours dispensé par le professeur (vous) pour réviser brièvement la matière du test qui sera soumis aux élèves (Chapitres UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré)). La récolte de données en tant que telle et la passation du test sont prévues lors de la deuxième heure. Selon le nombre de classes de troisième secondaire que vous avez, celles-ci seront réparties par le chercheur dans la situation 1 ou 2. Vous donnerez donc des consignes différentes selon que la classe soit dans telle ou telle situation. À la suite du test, les élèves seront invités à répondre à un questionnaire sur le concept de soi en mathématiques et sur les buts d'accomplissements. Enfin, il sera demandé aux élèves de signer un document de consentement.

2 Étape de révision

Cette révision poursuit plusieurs objectifs :

A) Rappeler la matière qui sera évaluée lors du test (Chapitre UAA3 et UAA4).

- Le contenu de cette heure de rappel est assez libre. Nous vous laissons le choix d'exposer les exercices, les théories, les techniques que vous souhaitez. Cela permet d'être en adéquation avec la façon dont vous avez introduit ces chapitres plus tôt dans l'année.

B) Annoncer l'évaluation

- Il se peut que les élèves soient surpris de revoir une matière travaillée plus tôt durant l'année ou d'interrompre un chapitre en cours. L'idée est de justifier cela par le fait qu'une évaluation type « Epreuve externe » est proposée à certains

enseignants de troisième en Fédération Wallonie-Bruxelles. Vous pouvez donner l'exemple des CEB, CE1D, etc. tout en précisant que cette évaluation n'est pas certificatives au contraire de ces deux épreuves.

Exemple : « *Nous allons interrompre très légèrement le chapitre que nous sommes en train de voir actuellement ensemble pour revenir sur le chapitre des fonctions que nous avons vu plus tôt cette année. Nous allons passer une heure à revoir ce chapitre puis vous aurez une évaluation lors d'une deuxième heure. Cette évaluation est proposée à d'autres classes de troisième de la Fédération Wallonie-Bruxelles. Les professeurs comme moi qui acceptent, la mettent en place dans leurs classes. Pour vous, cela ne change rien des évaluations traditionnelles que vous faites avec moi d'habitude. Elle sera notée et elle comptera pour le bulletin.* »

- À ce stade, il est important que les élèves ne sachent pas qu'il s'agit d'un projet de mémoire. Il leur sera communiqué après la passation du test à la fin de la deuxième heure. Ici, vous présentez l'évaluation comme n'importe quelle autre évaluation sommative réalisée durant l'année.

C) Communiquer la date de l'évaluation aux élèves.

- Les dates de l'heure de rappel et d'évaluation sont choisies par vos soins selon ce qui est le moins contraignant pour vous, pour votre programme et/ou pour ce que vous avez prévu en dehors de ce dispositif. Néanmoins, afin de ne pas engendrer un décalage temporel trop important, nous demandons que les deux heures soient séparées d'une semaine maximum.
- Dans le cas où les élèves seraient inquiets quant au délai que vous laissez entre ces deux heures, voici 3 arguments pour les rassurer :
 - a) Expliquer que ce test repose sur une matière vue précédemment durant l'année et qui est donc censée être maîtrisée.
 - b) Inviter ceux qui le souhaitent à s'entraîner avec les exercices vus plus tôt dans l'année.
 - c) Assurer que l'évaluation ne porte pas sur des éléments théoriques, mais sur du raisonnement.

3 Passation du test

Comme précisé précédemment, chacune de vos classes impliquées dans ce dispositif sera placée dans la situation 1 ou 2. Ce qui différencie ces deux situations est la façon de présenter l'évaluation. En effet, nous comparerons la performance des élèves en situation évaluative axée sur la performance (situation 1) à celle en situation d'évaluation axée sur la maîtrise (situation 2). Mis à part cela, les deux groupes passeront le même test et répondent au même questionnaire.

Pour que l'implémentation soit la plus semblable au sein de toutes les classes, nous vous proposons un script pour chacune des situations. Les consignes indiquées **en gras** dans le script ci-dessous doivent être **lues** telles quelles aux élèves, car il s'agit de phases fondamentales de l'expérimentation. Le reste ne doit pas obligatoirement être lu à la lettre.

3.1. Situation 1 (performance)

Temps consacré	Consignes
10'	<p><u>Introduction</u></p> <ul style="list-style-type: none">- Installation- Présentation évaluation <p>« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test permettra d'évaluer votre niveau dans cette matière. Ce test nous permettra également de comparer vos capacités à celles des autres étudiants de la classe. »</p> <ul style="list-style-type: none">- Consignes pour l'évaluation : <ul style="list-style-type: none">* Le carnet a été conçu de façon à ce que chaque élève puisse travailler seul. Les élèves n'ont donc, en principe, pas besoin d'explications supplémentaires.* Il est nécessaire que chaque élève dispose de sa calculatrice. L'utilisation du téléphone n'est pas autorisée. Insister pour que les élèves lisent attentivement les consignes et les questions et qu'ils examinent soigneusement les différentes représentations proposées (y compris les éventuels titres, titres des axes, etc.).* Certaines questions nécessitent un calcul, dont la réponse peut être un nombre décimal. Dans ce cas, on conseillera à l'élève d'arrondir sa réponse à deux chiffres après la virgule.

	<p>* Le vocabulaire du carnet a été soigneusement choisi et les concepteurs de l'épreuve considèrent que ce vocabulaire appartient au lexique de la discipline. Dans un souci de standardisation des conditions de passation, il est demandé de ne pas clarifier des mots incompris.</p> <ul style="list-style-type: none"> - Distribuez les feuilles. - Précisez qu'ils ont 30' pour réaliser ce test. - Invitez les élèves à d'abord compléter le questionnaire de la première page du test (questions personnelles).
30'	<p><u>Passation du test :</u></p> <ul style="list-style-type: none"> - Précisez aux élèves qu'ils ont trente minutes pour répondre aux questions d'introductions et aux questions de l'évaluation. - Ne répondez pas aux questions sur le contenu des items et ne fournissez aucune information particulière, réponse ou instruction sur quelque item que ce soit. - Si les questions des élèves concernent la manière de répondre (ex: l'endroit où indiquer sa réponse), vous pouvez intervenir. - Si un élève termine son test plus tôt, proposez-lui de relire et attendez que le temps soit écoulé. - Une fois que les trente minutes sont écoulées ou que tous les élèves ont terminé le test, reprenez les feuilles.
10'	<p><u>Passation Questionnaire :</u></p> <ul style="list-style-type: none"> - Distribuez les questionnaires et lire le texte d'introduction. <p><i>« Ce test est accompagné d'un petit questionnaire. Dans ce questionnaire, vous devez répondre à diverses questions pour lesquelles il n'y a pas de bonnes ou mauvaises réponses. Cette partie ne fait plus partie de l'évaluation. Je ne regarderai d'ailleurs pas vos réponses à ces questions. »</i></p> <p><u>Débriefing :</u></p> <ul style="list-style-type: none"> - Reprenez les questionnaires et expliquez le réel sens du test. <p><i>« En réalité, ce test est un peu particulier car, au-delà d'être un test habituel, normal, avec une influence sur votre bulletin, il s'inscrit également dans une recherche réalisée dans le cadre d'un mémoire. Cette expérimentation a pour but de voir comment la pression de l'évaluation influence vos performances. Lorsque le test aura été passé par tous les participants, les différents résultats seront observés afin de voir si certains liens peuvent être établis entre les résultats au test et les réponses fournies dans les questionnaires. »</i></p> <ul style="list-style-type: none"> - Les détails concernant les possibles répercussions sur les élèves redoublants sont dissimulés de manière à ne pas accroître la stigmatisation qu'ils pourraient subir de la part des autres élèves. - Distribuez les documents de consentement à faire signer par les élèves.

3.2. Situation 2 (maitrise)

Temps consacré	Consignes
10'	<p><u>Introduction</u></p> <ul style="list-style-type: none">- Installation- Présentation évaluation <p>« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test vous aidera à comprendre davantage ce que vous avez vu du chapitre sur les fonctions. Vous verrez que, même pendant le test, vous continuerez à apprendre et à développer des connaissances. »</p> <ul style="list-style-type: none">- Consignes pour l'évaluation :<ul style="list-style-type: none">* Le carnet a été conçu de façon à ce que chaque élève puisse travailler seul. Les élèves n'ont donc, en principe, pas besoin d'explications supplémentaires.* Il est nécessaire que chaque élève dispose de sa calculatrice. L'utilisation du téléphone n'est pas autorisée. Insister pour que les élèves lisent attentivement les consignes et les questions et qu'ils examinent soigneusement les différentes représentations proposées (y compris les éventuels titres, titres des axes, etc.).* Le vocabulaire du carnet a été soigneusement choisi et les concepteurs de l'épreuve considèrent que ce vocabulaire appartient au lexique de la discipline. Dans un souci de standardisation des conditions de passation, il est demandé de ne pas clarifier des mots incompris.* Certaines questions nécessitent un calcul, dont la réponse peut être un nombre décimal. Dans ce cas, on conseillera à l'élève d'arrondir sa réponse à deux chiffres après la virgule.- Distribuez les feuilles.- Précisez qu'ils ont 30' pour réaliser ce test.- Invitez les élèves à d'abord compléter le questionnaire de la première page du test (questions personnelles).

30'	<p><u>Passation du test :</u></p> <ul style="list-style-type: none"> - Précisez aux élèves qu'ils ont trente minutes pour répondre aux questions d'introductions et aux questions de l'évaluation. - Ne répondez pas aux questions sur le contenu des items et ne fournissez aucune information particulière, réponse ou instruction sur quelque item que ce soit. - Si les questions des élèves concernent la manière de répondre (ex: l'endroit où indiquer sa réponse), vous pouvez intervenir. - Si un élève termine son test plus tôt, proposez-lui de relire et attendez que le temps soit écoulé. - Une fois que les trente minutes sont écoulées ou que tous les élèves ont terminé le test, reprenez les feuilles.
10'	<p><u>Passation Questionnaire :</u></p> <ul style="list-style-type: none"> - Distribuez les questionnaires et lire le texte d'introduction. <p><i>« Ce test est accompagné d'un petit questionnaire. Dans ce questionnaire, vous devez répondre à diverses questions pour lesquelles il n'y a pas de bonnes ou mauvaises réponses. Cette partie ne fait plus partie de l'évaluation. Je ne regarderai d'ailleurs pas vos réponses à ces questions. »</i></p> <p><u>Débriefing :</u></p> <ul style="list-style-type: none"> - Reprenez les questionnaires et expliquez le réel sens du test. <p><i>« En réalité, ce test est un peu particulier car, au-delà d'être un test habituel, normal, avec une influence sur votre bulletin, il s'inscrit également dans une recherche réalisée dans le cadre d'un mémoire. Cette expérimentation a pour but de voir comment la pression de l'évaluation influence vos performances. Lorsque le test aura été passé par tous les participants, les différents résultats seront observés afin de voir si certains liens peuvent être établis entre les résultats au test et les réponses fournies dans les questionnaires. »</i></p> <ul style="list-style-type: none"> - Les détails concernant les possibles répercussions sur les élèves redoublants sont dissimulés de manière à ne pas accroître la stigmatisation qu'ils pourraient subir de la part des autres élèves. - Distribuez les documents de consentement à faire signer par les élèves.

4 Après le test

4.1. Correction et résultats

Une fois les tests passés, le chercheur prendra contact avec vous afin de récupérer les différents documents de vos classes. La correction des tests sera prise en charge par le chercheur. Lorsque celle-ci sera achevée, vous récupérerez les évaluations corrigées afin de les remettre aux élèves. Vous pouvez, si vous le souhaitez, corriger les tests avec l'ensemble de la classe afin d'apporter un retour constructif à cette épreuve.

Divers éléments vous seront communiqués après la passation et la correction des tests. Tout d'abord, vous obtiendrez le résultat total au test de chaque élève. Ensuite, vous serez également transmis les résultats au test de chaque élève pour chaque compétence (Connaître / Appliquer / Transférer).

En outre, il vous sera possible d'obtenir le taux de réussite de vos classes au test, le taux de réussite pour chaque item, etc. Cela vous donnera l'occasion de voir le niveau de vos élèves dans cette matière dans une épreuve externe. Enfin, nous vous enverrons par mail les résultats généraux de l'étude.

4.2. Conséquences du test sur le bulletin

Comme expliqué précédemment, ce test est une évaluation sommative ce qui signifie qu'elle a une influence sur les points du bulletin des élèves qui participent à cette expérimentation. Nous encourageons très fortement les enseignants pour qu'ils prennent réellement en compte cet aspect du test, car nous souhaitons éviter que des élèves étudient et s'investissent pour une épreuve qui finalement ne « compterait » pas. Cela pourrait engendrer de la frustration chez certains élèves.

Cependant, nous vous conseillons vivement d'accorder une influence dans le bulletin relativement faible au vu des conditions « particulières » de cette évaluation. En effet, nous faisons tout de même l'hypothèse que les performances des élèves dans la condition 1 (axée sur la performance) seront plus faibles. Nous faisons également l'hypothèse que certaines catégories d'élèves (doublants et filles notamment) pourraient être en difficulté par rapport à d'autres selon les conditions. Ainsi, accorder une trop grosse influence à cette évaluation en sachant cela ne serait pas juste.

En conclusion, même si cette évaluation est notée, s'il s'avère qu'une des deux modalités a dans l'ensemble un score global significativement plus faible, alors il vous sera demandé d'appliquer un coefficient correctif afin de prendre en compte le désavantage vécu dans cette modalité. De cette manière nous souhaitons réduire tous les désagréments possibles et apporter, quelles que soient les conditions, un apport bénéfique à tous les élèves.

**Nous vous remercions à nouveau vivement pour votre précieuse participation
à la gestion et à la mise en place de cette étude.**

Si vous avez des questions ou des incertitudes à l'issue de cette lecture ou pendant l'administration de test, n'hésitez pas à me contacter. Voici mes coordonnées :

Flore Nicolas

**Étudiant en Sciences de l'Éducation
Université de Liège
Rue Thier-Hamal 8, 4631 Soumagne**

**Tél. : 0470/ 33 30 16
Nicolas.Flore@student.uliege.be**

Annexe 2 : Points de matière pour la période de rappel

Rappel

Voici quelques notions théoriques qu'il peut être utile d'aborder lors de cette heure de révision. Ces différentes notions peuvent idéalement être accompagnées de quelques exercices.

UAA3 Approche graphique des fonctions :

- VD-VI
- Image $f(x) = y$
- Domaine et Ensemble Image
- Racines et Ordonnées à l'origine
- Tableau de signes et de variations

UAA4 Fonctions du premier degré :

- Tableau – Graphique – Formule
- $y = mx + p$
- Affine – Linéaire – Constante
- Croissante – Décroissante
- $m = DV / DH = y_A - y_B / x_A - x_B$
- Déterminer l'équation d'une droite

ÉVALUATION SOMMATIVE 2024
MATHÉMATIQUES

NOM :

PRÉNOM :

CLASSE :

N° D'ORDRE :

ÉCOLE :

Avant d'aller plus loin, veuillez répondre aux questions suivantes :

* *Quel âge avez-vous ?*

* *Êtes-vous ...*

- ☐ Un homme
- ☐ Une femme
- ☐ Je ne préfère pas le dire

* *Quels cours avez-vous en « option » ?*

* *Combien d'heures de mathématiques avez-vous par semaine ?*

* *Avez-vous déjà redoublé durant votre scolarité (de la maternelle jusqu'à aujourd'hui) ? Si oui, combien de fois ?*

- ☐ Non
- ☐ Oui (.... fois)

ATTENTION

Tu peux annoter les pages, réaliser des schémas ou mettre en évidence des éléments importants.

Il est inutile d'effacer les démarches faites au brouillon.

Dans les pages qui suivent, certaines expressions peuvent nécessiter une précision :

- un « zéro d'une fonction » et une « racine d'une fonction » sont employés comme des synonymes ;
- l'expression analytique d'une fonction est une écriture mathématique formulée, par exemple comme ceci :

- $f(x) = 3x + 1$

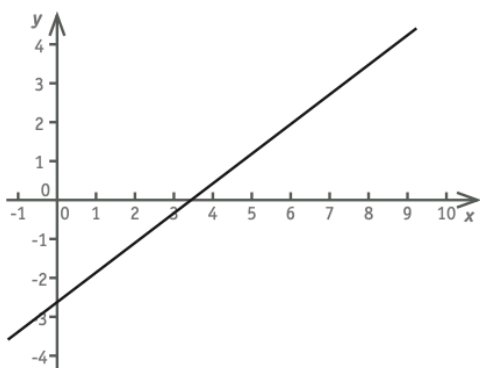
- $y = mx + p$

- $y = -4x$

QUESTION 1

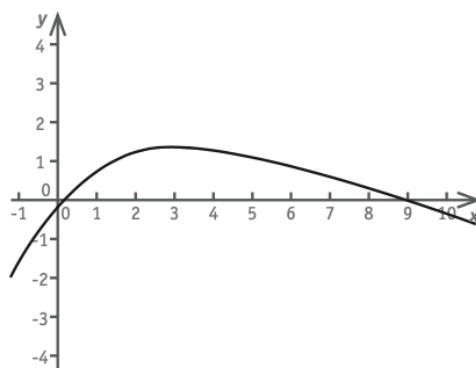
Voici des représentations graphiques de relations entre deux grandeurs. Identifie toutes celles qui sont des représentations de fonctions.

ENTOURE OUI lorsqu'il s'agit d'une fonction et NON dans le cas contraire.



Cette relation est-elle une fonction ?

OUI | NON

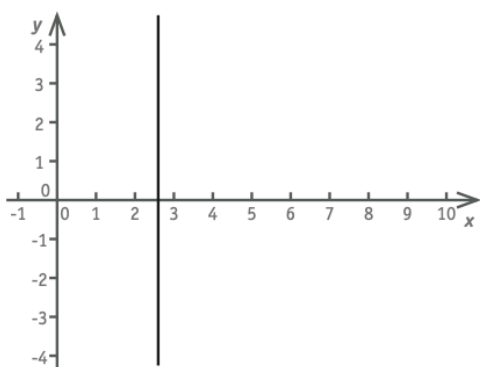


Cette relation est-elle une fonction ?

OUI | NON

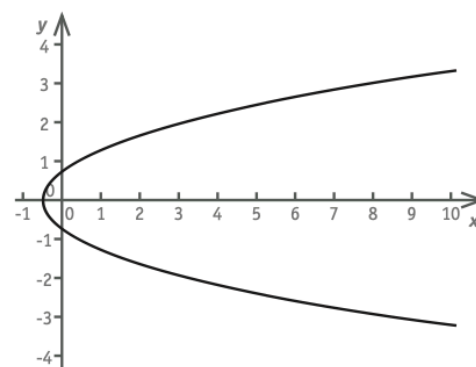
☐ 1a

☐ 1b



Cette relation est-elle une fonction ?

OUI | NON



Cette relation est-elle une fonction ?

OUI | NON

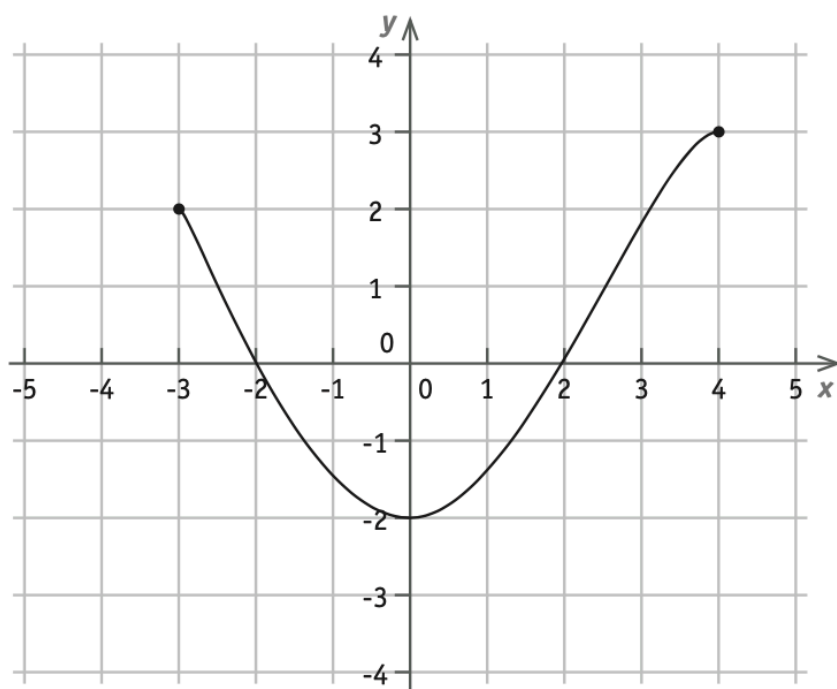
☐ 1c

☐ 1d

QUESTION 2

Quel est le domaine de définition de chaque fonction représentée graphiquement ?

COCHE la case adéquate pour chacune des deux fonctions proposées.



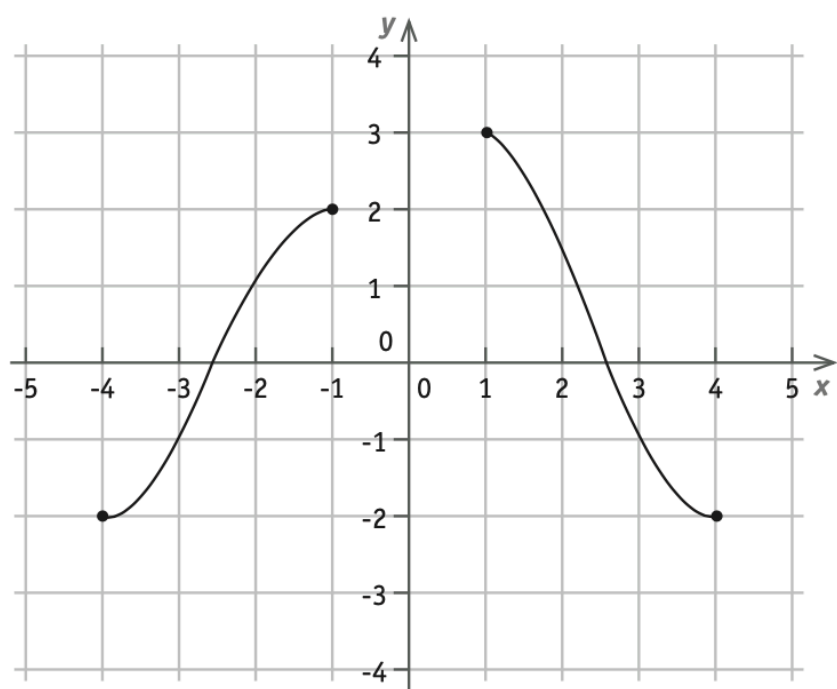
☐ $[-2 ; 3]$

☐ $[3 ; -2]$

☐ $[-3 ; 4]$

☐ $[4 ; -3]$

☐ 2a



☐ $[-4 ; -1] \cup [1 ; 4]$

☐ $[-4 ; 4]$

☐ $[-4 ; 0[\cup]0 ; 4]$

☐ $[-2 ; 3]$

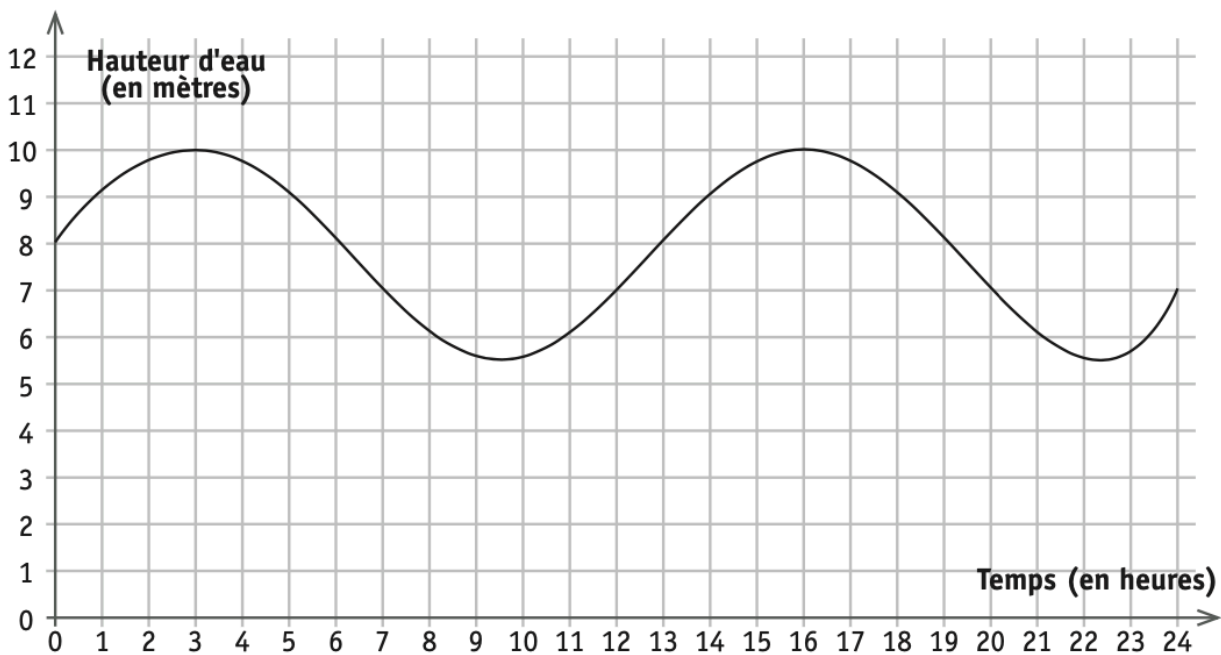
☐ 2b

QUESTION 3

Dans sa zone d'estuaire, un fleuve subit l'effet des marées maritimes. En effet, on constate que la profondeur de l'eau y varie selon la montée et la descente du niveau de la mer.

La Tamise, en Angleterre, est un fleuve dont les marées fluviales ont une grande amplitude.

Le graphique ci-dessous représente l'évolution de la profondeur de la Tamise le 5 mai 2017.



Sur base du graphique, **RÉPOND**S aux questions suivantes.

- a) Quelle différence maximale de hauteur entre une marée haute et une marée basse de la Tamise a-t-on observée le 5 mai 2017 ?

3a

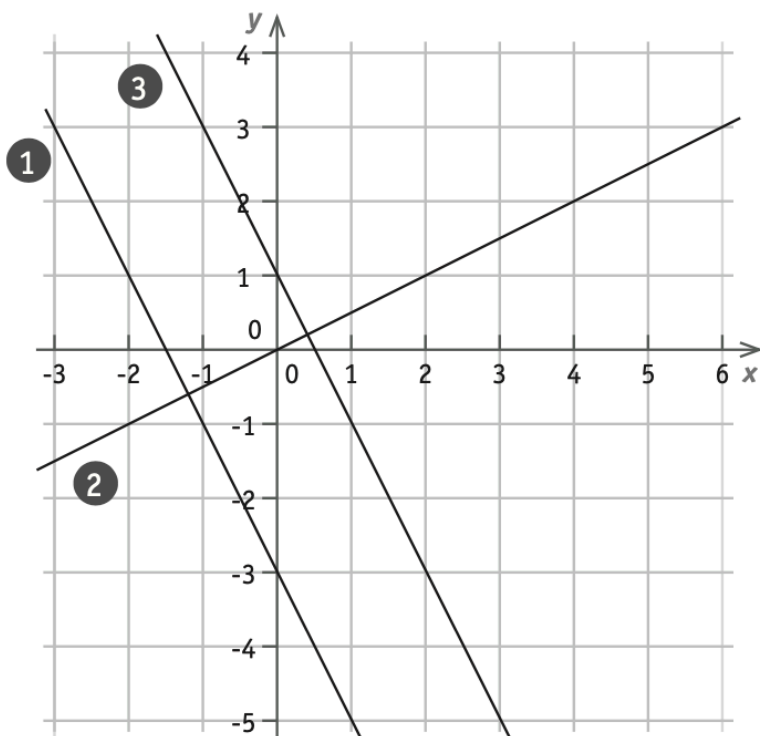
- b) Sachant que pour naviguer en sécurité, un bateau a besoin de minimum 7 mètres de profondeur d'eau, quelles sont les périodes du 5 mai 2017 durant lesquelles la navigation était sécurisée sur la Tamise ?

3b

QUESTION 4

Voici des représentations graphiques et des expressions analytiques de fonctions du premier degré.

Représentations graphiques



Expressions analytiques

$$f(x) = -2x + 1$$

$$g(x) = 2x - 3$$

$$h(x) = -3$$

$$i(x) = -x + 1$$

$$j(x) = -2x - 3$$

$$k(x) = 0,5x$$

Associe chaque représentation graphique à son expression analytique.

COMPLÈTE les phrases suivantes :

La représentation graphique ① correspond à la fonction _____

4a

La représentation graphique ② correspond à la fonction _____

4b

La représentation graphique ③ correspond à la fonction _____

4c

QUESTION 5

DÉTERMINE l'expression analytique d'une fonction f du premier degré dont le graphique passe par les points A et B de coordonnées $(1 ; 2)$ et $(3 ; 8)$.

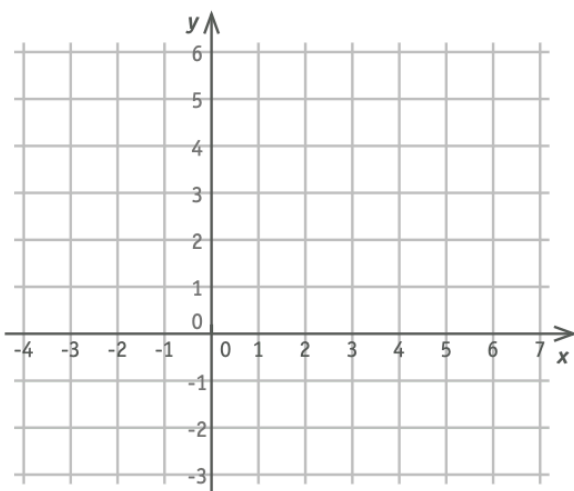
5

L'expression analytique de la fonction f est : $f(x) =$ _____

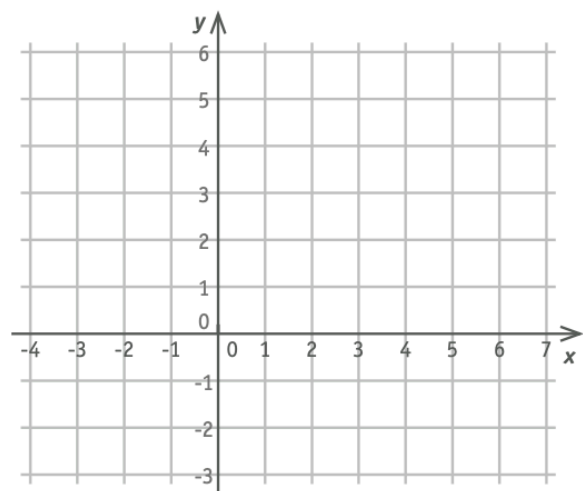
QUESTION 6

Trace les graphiques des fonctions données dans les repères ci-dessous.

$$f(x) = 3$$



$$g(x) = -2x + 1$$



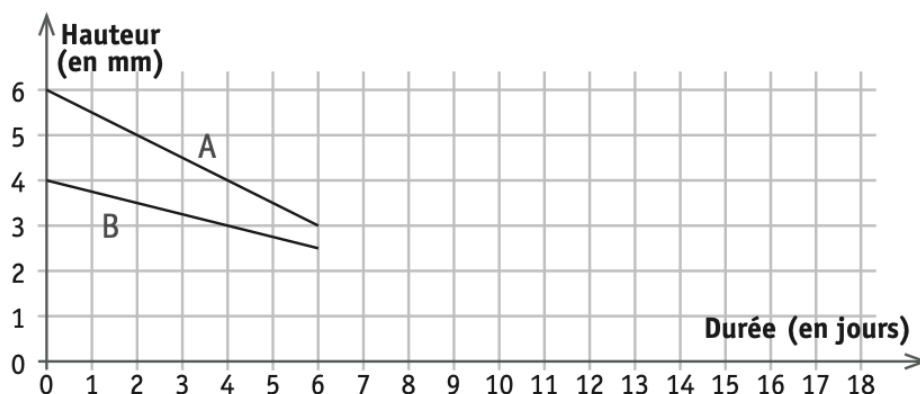
6a

6b

QUESTION 7

Deux éprouvettes A et B contiennent des huiles différentes qui s'évaporent au fil des jours.

Le graphique ci-dessous modélise la hauteur (en millimètres) de l'huile restant dans les éprouvettes en fonction du nombre de jours écoulés. Le graphique représente la situation des six premiers jours.



RÉPONDS aux questions suivantes. Tu peux dessiner sur le graphique.

- D'après le graphique, après combien de jours l'évaporation de l'huile contenue dans l'éprouvette A sera-t-elle complète ?

7a

- D'après le graphique, après combien de jours les quantités d'huile contenues dans les deux éprouvettes seront-elles identiques ?

7b

Annexe 4 : Questionnaire contextuel

QUESTIONNAIRE – 3eG – 2024
MATHÉMATIQUES

NOM :

PRÉNOM :

CLASSE :

N° D'ORDRE :

ÉCOLE :

Bonjour,

Tu trouveras dans ce dossier une série de questions que nous te demande de lire attentivement afin d'y répondre le plus précisément possible.

Il n'y a pas de « bonnes » ou de « mauvaises » réponses. Réponds simplement en indiquant sincèrement ce qui correspond à ta situation.

Dans tous les cas, il s'agit de répondre en cochant la case qui te convient. Si tu as fait une erreur en cochant une case, barre la mauvaise case (ou efface proprement ta croix), puis coche la bonne case.

Tes réponses seront combinées avec celles d'autres élèves pour donner des résultats totaux, où aucun élève particulier ne pourra être identifié. Tout ce que tu indiqueras dans ce questionnaire restera confidentiel.

Nous te remercions vivement pour ta collaboration !

Section 1 :

Dans quelle mesure es-tu d'accord avec les phrases ci-dessous ? (ne coche qu'une seule case par ligne).

		Pas du tout d'accord	Pas d'accord	D'accord	Tout à fait d'accord
1)	Les mathématiques sont l'une de mes matières préférées.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
2)	J'ai des facilités en mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
3)	Je veux avoir de bons résultats en mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
4)	J'ai peur d'avoir de mauvaises notes en mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
5)	J'ai peur d'échouer en mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
6)	J'ai souvent peur d'avoir des difficultés en cours de mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
7)	Je me sens perdu quand je travaille sur un problème mathématique.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
8)	Je deviens très nerveux(se) quand je travaille sur des problèmes de mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
9)	Je suis très tendu(e) quand je dois faire des devoirs de mathématiques à la maison.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

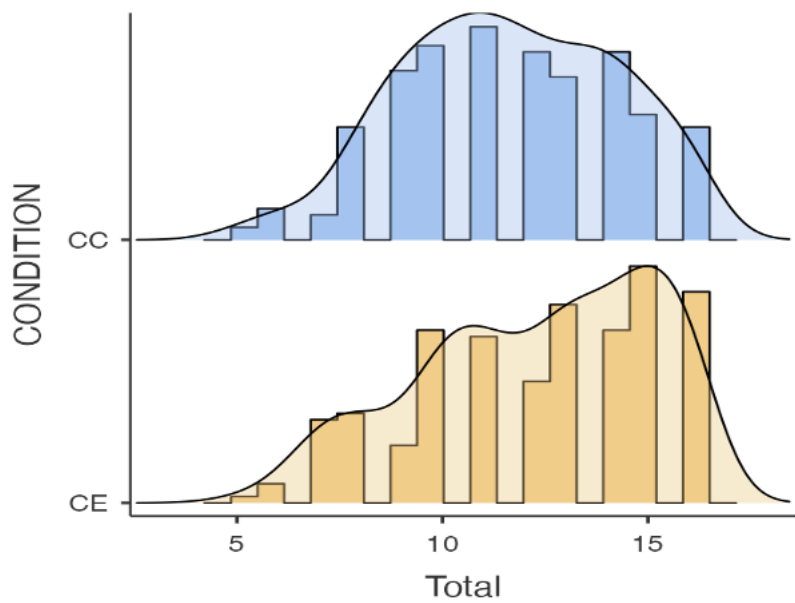
Section 2 :

Par rapport au cours de mathématiques :

Dans quelle mesure es-tu d'accord avec les phrases ci-dessous ? (ne coche qu'une seule case par ligne).

		Pas du tout d'accord	Pas d'accord	D'accord	Tout à fait d'accord
1)	Il est important pour moi de mieux réussir que les autres.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
2)	Je m'inquiète de ne pas apprendre autant que je le pourrais dans ce cours.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
3)	Mon but dans ce test et dans ce cours est d'avoir de meilleures notes que la plupart des étudiants.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
4)	Parfois, j'ai peur de ne pas comprendre le contenu de ce cours de manière aussi approfondie que je le souhaiterais.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
5)	Ma peur d'échouer dans ce cours est souvent ce qui me motive.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
6)	Il est important pour moi de bien réussir comparativement aux autres dans ce cours et à ce test.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
7)	C'est important pour moi de comprendre de façon aussi approfondie que possible les contenus du cours de mathématiques.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
8)	Mon but en mathématiques est d'éviter de mal réussir.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
9)	En mathématiques, je veux apprendre autant que possible.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
10)	Je désire maîtriser complètement les contenus de ce cours.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
11)	Je veux seulement éviter d'échouer dans ce cours.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
12)	Je suis parfois soucieux du fait que je pourrais ne pas apprendre tout ce qu'il y a à apprendre dans ce cours.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

Annexe 5 : Histogramme de densité des deux échantillons (deux conditions)



Concerne : recrutement pour une recherche de mémoire sur les éventuelles influences des contextes évaluatifs sur les performances des élèves selon leur parcours scolaire.

Madame, Monsieur,

Je suis actuellement en deuxième année du Master de sciences de l'Éducation à l'Université de Liège. De ce fait, dans le but de finaliser ma formation universitaire, je me permets de vous écrire cette lettre afin de vous demander l'autorisation de soumettre un test de mathématique sommatif aux élèves de troisième secondaire de votre école, dans le cadre de mon mémoire s'intitulant : « Le contexte évaluatif influence-t-il de façon différentielle la performance des élèves en fonction de leur parcours scolaire ? ».

Mon étude a pour objectif, au travers d'un test de mathématique noté et ayant une influence dans le bulletin, de comparer les performances de diverses catégories d'élèves de troisième secondaire de l'enseignement général de la FWB dans deux contextes évaluatifs différents (l'un axé sur la performance et l'autre sur la maîtrise).

Pour mon étude, j'aimerais recruter environ 400 élèves de troisième secondaire de l'enseignement général de la Fédération Wallonie Bruxelles. Il leur sera demandé, s'ils acceptent de participer à ma recherche, de répondre à un test de mathématique ainsi que de répondre à deux questionnaires. Le premier questionnaire porte sur les buts poursuivis par les élèves durant le test (ex : « Durant ce test il était important pour moi de mieux réussir que les autres élèves de la classe »). Le deuxième questionnaire s'intéresse quant à lui au sentiment d'efficacité personnelle ressenti par les élèves durant le test.

Ma recherche a fait l'objet d'un avis positif de la Commission de Vigilance éthique du Département des Sciences de l'Éducation en date du 19/02/2024 et toutes les précautions nécessaires seront prises afin de garantir l'anonymat et le respect de la vie privée des participants.

1) Projet

Ce projet de recherche a pour but d'observer les effets du contexte évaluatif d'une évaluation sommative à enjeu en mathématique sur les performances des élèves en retard en troisième secondaire. Notre choix de s'intéresser particulièrement aux élèves qui ont connu le redoublement est motivé par plusieurs facteurs. Nous nous sommes rendu compte que peu d'études se concentraient sur cette population. Tout d'abord, car un certain nombre de recherches se sont centrées sur des groupes ethniques, sur les genres, sur les statuts socio-économiques, mais assez peu sur les redoublants. De plus, l'enseignement de la Fédération Wallonie-Bruxelles compte l'un des taux d'élèves en retard les plus importants parmi les pays de l'OCDE.

Concrètement, les élèves passeront un test de mathématiques qui sera noté et que vous pourrez utiliser comme une évaluation traditionnelle qui compte pour la note finale du bulletin. Les items de ce test portent sur les chapitres UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré). L'idée est de manipuler le contexte évaluatif de deux manières différentes afin de voir si certaines manières de présenter l'évaluation peuvent être davantage bénéfiques ou préjudiciables pour les élèves en retard scolaire. Nous comparerons la performance des élèves en situation évaluative axée sur la performance (situation 1) à celle en situation d'évaluation axée sur la maîtrise (situation 2). Les deux groupes passeront le même test.

2) Déroulement

Si vous acceptez de participer à cette recherche, l'expérimentation nécessitera deux heures de cours pour chaque classe. La première heure consiste en un rappel, un cours dispensé par le professeur (vous) pour réviser brièvement la matière du test qui sera soumis aux élèves (Chapitres UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré)). De cette manière, le contenu du test est rappelé et les élèves ne sont pas interrogés sur un sujet qu'ils ne connaissent pas ou qu'ils pourraient avoir oublié. La récolte de données en tant que telle et la passation du test sont prévues lors de la deuxième heure. Selon le nombre de classes de troisième secondaire que vous avez, celles-ci seront réparties par le chercheur dans la situation 1 ou 2. Tous les élèves d'une même classe seront dans la même situation. En tant que professeur, vous donnerez donc des consignes différentes selon que la classe soit dans telle ou telle situation. Afin de ne pas engendrer un décalage temporel trop important, nous demandons que les deux heures soient séparées d'une semaine maximum.

3) Consignes de passation

Pour la deuxième heure, voici les consignes et les points d'attention à communiquer aux élèves.

A. Présentation du test

Consigne commune pour les deux situations :

« Il y a quelque temps OU l'heure passée, nous avons fait une heure de révision sur les fonctions. Comme annoncé, nous allons aujourd'hui réaliser un test qui sera noté. Sachez que le test que vous allez recevoir est en réalité mis en place dans d'autres écoles et d'autres classes de troisième. C'est un peu comme une Évaluation Externe (ex : CEB, CE1D...). Cela ne change rien pour vous. Seulement, la première chose à faire est de remplir le petit questionnaire contextuel qui nécessite quelques courtes informations personnelles. Vous le retrouverez sur les premières pages. »

Consigne pour la situation 1 :

« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test permettra d'évaluer votre niveau dans cette matière. Ce test nous permettra également de comparer vos capacités à celles des autres étudiants de la classe. »

Consigne pour la situation 2 :

« Sur base de ce test, vous recevrez une note qui interviendra dans votre note de période. Ce test vous aidera à comprendre davantage ce que vous avez vu du chapitre sur les fonctions. Vous verrez que, même pendant le test, vous continuerez à apprendre. »

B. Questionnaires sur les buts d'accomplissement et le sentiment d'efficacité personnelle

Consigne commune pour les deux conditions :

« Ce test est accompagné de deux petits questionnaires. Dans ces questionnaires, vous devez répondre à diverses questions pour lesquelles il n'y a pas de bonnes ou mauvaises réponses. Donc, une fois que vous avez terminé votre test de mathématique, vous levez la main et je passerai alors vous donner ces courts questionnaires que vous complétez. »

C. Debriefing

Lorsque que tous les élèves ont rendu le test et les deux questionnaires ou que le temps est écoulé, vous pouvez leur annoncer le véritable but de ce test afin de les rassurer.

« En réalité, ce test est un peu particulier car, au-delà d'être un test habituel, normal, avec une influence sur votre bulletin, il s'inscrit également dans une recherche dans le cadre d'un mémoire. Cette expérimentation a pour but de voir comment la pression de l'évaluation influence vos performances. Lorsque le test aura été passé par tous les participants, les différents résultats seront observés afin de voir si certains liens peuvent être établis entre les résultats au test et les réponses fournies dans les questionnaires. »

4) Informations complémentaires

Afin de finaliser votre participation, veuillez répondre aux questions suivantes :

- De combien de classes de troisième secondaire êtes-vous l'enseignant(e) ?

.....

- Vos classes rassemblent-elles des élèves issus d'une option en particulier ? Si oui, laquelle ?

.....

- > Est-ce que les chapitres sur les fonctions UAA3 (Approche graphique d'une fonction) et UAA4 (Le premier degré) ont déjà été vus dans toutes vos classes de troisième secondaire ?

☐ Oui ☐ Non

- Si oui, à quelle période de l'année et combien d'heures de cours (approximativement) y avez-vous consacrées ?

-
- Je souhaite prendre part à cette recherche et accepte de mener le dispositif détaillé ci-dessus dans mes classes de mathématique de troisième secondaire :
- ☐ Oui ☐ Non

Si oui, merci de renvoyer ce document par courriel l'adresse mail suivante :
Nicolas.Flore@student.uliege.be

NOM Prénom :

Signature :

Toutes les questions que vous vous posez sur ce travail peuvent m'être adressées directement par mail à Nicolas.Flore@student.uliege.be ou par téléphone au 0470/33.30.16.

En vous remerciant pour l'intérêt que vous porterez à mon travail, et en restant à votre entière disposition, je vous prie d'agréer, Madame, Monsieur, l'expression de mes salutations respectueuses.

Flore Nicolas,
Étudiant du Master en Sciences de l'Éducation
