Aix*Marseille université
Socialement engagée

NNT : 2024AIXM0216

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 08 juillet 2024 par

Marjorie ARMANDO

Améliorer les performances des enfants en mathématiques :
un agent virtuel pédagogique pour lutter contre les effets de menace du stéréotype

Discipline
Informatique

École doctorale
ED 184 Mathématiques et Informatique

Laboratoires/Partenaires de recherche
Laboratoire d'Informatique & Systèmes UMR 7020 CNRS / AMU / UTLN

Centre de Recherche en Psychologie et Neurosciences UMR 7077 CNRS / AMU

Pôle pilote Ampiric, Institut National Supérieur du Professorat et de l'Éducation, Aix-Marseille Université, Marseille, France

Composition du jury

Justine CASSELL Professeure INRIA Paris	Rapportrice
Pascal PANSU Professeur Université de Grenoble	Rapporteur
Jean-Claude MARTIN Professeur Université Paris Saclay	Examinateur
Agnès BLAYE Professeure Aix-Marseille Université	Présidente du jury
Magalie OCHS Maîtresse de conférences Aix-Marseille Université	Directrice de thèse
Isabelle RÉGNER Professeure Aix-Marseille Université	Co-directrice de thèse

LIS CRPN cnrs ILCB Ampiric Pôle pilote de formation des enseignants et de recherche pour l'éducation

Affidavit

Je soussigné, Marjorie Armando, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Magalie Ochs et d'Isabelle Régner, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille, le 21 avril 2024



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Liste de publications et participations aux congrès et conférences

Article de revue internationale avec comité de lecture :

Armando, M., Ochs, M., & Régner, I. (2022). The Impact of Pedagogical Agents' Gender on Academic Learning : A Systematic Review. *Frontiers in Artificial Intelligence*, 5, 862997. https://doi.org/10.3389/frai.2022.862997

Article de workshop international avec comité de lecture :

Armando, M., Régner, I., & Ochs, M. (2023). Toward a Tool Against Stereotype Threat in Math : Children's Perceptions of Virtual Role Models. *International Workshop on Multimodal, Interactive Interfaces for Education in International Conference on Multimodal Interaction (ICMI)*, Paris, France. https://doi.org/10.1145/3610661.3616179

Article de workshop national avec comité de lecture :

Armando, M., Régner, I., & Ochs, M. (2021). Les effets du genre des agents virtuels sur l'apprentissage. *Workshop sur les "Affects, Compagnons Artificiels et Interactions" (WACAI)*, Île d'Oléron, France. https://hal.science/hal-03366193v2

Participations aux congrès et écoles d'été au cours de la période de thèse :

1. 14 novembre 2023 - Journée Scientifique & Restitution des Ressources d'Ampiric, Marseille : « Améliorer les performances des enfants en maths : un agent pédagogique virtuel pour lutter contre les effets de la Menace du Stéréotype. ». Présentation orale.
2. 13 octobre 2023 - ICMI 2023 International Workshop on Multimodal, Interactive Interfaces for Education, Paris : « Toward a Tool Against Stereotype Threat in Math : Children's Perceptions of Virtual Role Models. ». Présentation orale.
3. 05 octobre 2023 - Colloque international en hommage à Jacques Ginestié, Marseille : « Améliorer les performances des enfants en maths : un agent pédagogique virtuel pour lutter contre les effets de la Menace du Stéréotype. ». Présentation orale.

4. 04 juillet 2023 - Plate-Forme Intelligence Artificielle (PFIA) 2023, Strasbourg : « Vers un outil pour contrer la Menace du Stéréotype : études pilotes sur les perceptions des enfants à l'égard des agents virtuels. ». Présentation orale.
5. 12-16 juin 2023 - École d'été en Traitement Automatique des Langues (ETAL), Marseille.
6. 30 juin 2022 - Plate-Forme Intelligence Artificielle (PFIA) 2022, Saint-Etienne : « Moins féminine donc plus compétente ? Étude perceptive de l'influence de la féminité et masculinité des personnages virtuels sur la perception des utilisateurs. ». Présentation orale.
7. 09-10 juin 2022 - Journées Scientifiques du LIS (JSL), Carry-le-Rouet : « Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype ». Présentation d'un poster, démonstration des personnages virtuels.
8. 08 juin 2022 - Journée Scientifique & Restitution des Ressources d'Ampiric, Marseille : « Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype ». Présentation d'un poster.
9. 14 octobre 2021 - Workshop sur les Affects, Compagnons Artificiels et Interactions (WACAI), Ile d'Oléron : « Les effets du genre des agents virtuels sur l'apprentissage. ». Présentation orale.
10. 24 août 2021 – Journée d'études « Femmes, genre et numérique : où est le problème ? », Namur, Belgique : « Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype ». Présentation orale.
11. 12 mars 2021 – Journée Annuelle des Doctorants, en ligne : « Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype ». Présentation orale.

Diffusions de l'information à un public non spécialiste au cours de la période de thèse :

1. 17-21 avril 2023 - Les Cigales au CIRM, stage de mathématiques pour lycéennes, Marseille. Encadrante de 2 groupes de lycéennes.
2. 06 mars 2023 – Journée « Femmes et sciences », Lycée Marseilleveyre, Marseille : « Présentation des métiers de la recherche à travers ma thèse ». Présentation orale et démonstration des personnages virtuels.
3. 08 décembre 2022 - Lycée de Tour Sainte, Marseille : « Présentation des métiers de la recherche à travers ma thèse ». Présentation orale et démonstration des personnages virtuels.
4. 02 novembre 2022 - Les Cigales au CIRM, stage de mathématiques pour lycéennes, Marseille : « Présentation des métiers de la recherche à travers ma thèse ». Présentation orale et démonstration des personnages virtuels.

5. 18 octobre 2022 - Collège Roumanille, Avignon : « Résultats préliminaires de la thèse : Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype. ». Présence de Christophe Mazuyer (IA-IPR de mathématiques) et du personnel. Présentation orale.
6. 20 septembre 2022 – The Conversations : « Réussite en maths : lutter contre les stéréotypes de genre avec le numérique. ». Article.
7. 17 mai 2022 – Lancement du laboratoire de maths du collège Roumanille, Avignon : « Améliorer les performances des enfants en maths : un compagnon virtuel pour lutter contre les effets de la Menace du Stéréotype ». Présence du personnel, de Claudie François-Gallin (Inspectrice d'académie - Directrice académique des services de l'Education nationale de Vaucluse), de Vincent Valéry (Directeur de l'Ecole académique de la formation continue), d'Odile Aubert (référente de l'éducation prioritaire), et de Bernard Beignier (recteur de l'académie Aix-Marseille). Présentation orale.
8. 07 mars 2022 – Journée « Femmes et sciences », Lycée Marseilleveyre, Marseille : « Présentation des métiers de la recherche à travers ma thèse ». Présentation orale et démonstration des personnages virtuels.

Résumé

De nombreuses recherches en cognition sociale ont montré que les stéréotypes de genre, qui véhiculent l'idée d'une infériorité intrinsèque des filles en mathématiques, peuvent nuire à leurs performances aux tests standardisés. Ce phénomène, connu sous le nom de menace du stéréotype, peut être réduit notamment en présentant des modèles féminins de réussite en sciences qui expliquent leurs bons résultats par le travail. Dans le contexte de la place grandissante du numérique éducatif, cette thèse se situe à l'interface de l'informatique et de la psychologie sociale, et poursuit un double objectif. Premièrement, il s'agit de développer des agents virtuels pédagogiques en 3D, adaptés à l'âge et au genre des élèves, capables d'interactions verbales et non-verbales pour créer un lien avec les élèves appelé "rapport virtuel" et pour expliquer leur réussite en mathématiques par leur travail et leurs efforts. Deuxièmement, il s'agit de tester si ces agents virtuels peuvent réduire l'effet de menace du stéréotype sur les performances des collégiennes en mathématiques. Après un travail informatique de modélisation des agents interactifs et de création d'un environnement virtuel d'apprentissage en mathématiques, deux études pilotes réalisées auprès de 144 élèves ont permis de vérifier que filles et garçons perçoivent les agents virtuels comme des modèles de réussite en mathématiques, s'identifient à eux, et cela est amplifié par le rapport virtuel instauré par les agents. Dans l'étude principale de la thèse, des élèves de 6e et 5e étaient guidés par un agent virtuel dans leur révision de notions mathématiques juste avant de passer un test sur ces notions. L'agent virtuel était soit féminin soit masculin, et variait dans sa façon d'expliquer sa réussite en mathématiques (absence d'explication *versus* explication centrée sur le travail et les efforts *versus* explication centrée sur le travail et les efforts avec création d'un rapport virtuel). Les résultats indiquent que l'agent féminin instaurant un rapport virtuel permet à tous les élèves d'obtenir de meilleures notes en mathématiques. Tout en étant cohérents avec la littérature sur les modèles de réussite en sciences, nos résultats contribuent à identifier, pour la première fois, leurs conditions d'efficacité dans le cadre de l'utilisation du numérique éducatif. Nous proposons plusieurs outils pour les chercheurs et les enseignants, notamment une version française du Test des Associations Implicites genre-mathématiques-lecture pour enfants/pré-adolescents, et un prototype d'application avec un agent virtuel pédagogique, où l'on peut adapter la leçon et les exercices.

Mots clés : agent virtuel pédagogique, menace du stéréotype, rapport virtuel, numérique éducatif, stéréotype de genre

Abstract

Many studies in social cognition have shown that gender stereotypes, which convey the idea of intrinsic inferiority of girls in mathematics, can impair their performance on standardized tests. This phenomenon, known as stereotype threat, can be reduced by presenting female role models in sciences who explain their success with hard work. In the context of the growing importance of educational technology, this thesis is at the intersection of computer science and social psychology, with two objectives. Firstly, we aim to develop 3D virtual pedagogical agents, adapted to students' age and gender, capable of verbal and non-verbal interactions to establish a connection with students called "virtual rapport", and capable to explain their success in mathematics through their work and efforts. Secondly, we seek to test whether these virtual agents can reduce stereotype threat effects on middle school girls' mathematics performance. After creating a computationnal model of interactive agents and create a virtual mathematics learning environment, two pilot studies involving 144 students confirmed that both boys and girls perceive virtual agents as successful mathematics role models, identify to them, and this is amplified by the virtual rapport established by the agents. In the thesis' main study, 6th and 5th graders in France were guided by a virtual agent in their revision of mathematical concepts just before taking a test on those concepts. The virtual agent was either female or male and varied in how it explained its success in mathematics (no explanation versus explanation focused on work and effort versus explanation focused on work and effort with building virtual rapport). The results indicate that the female agent establishing virtual rapport enables all students to achieve better grades in mathematics. While consistent with the literature on successful role models in sciences, our results contribute to identify, for the first time, their conditions of effectiveness in the context of educational technology. We propose several tools for researchers and teachers, including a French version of the Implicit Association Test gender-mathematics-reading for children/pre-adolescents, and a prototype application with a virtual pedagogical agent, where the lesson and exercises can be modified.

Keywords: pedagogical virtual agent, stereotype threat, virtual rapport, education technology, gender stereotype

Remerciements

Je tiens tout d'abord à remercier mes directrices de thèse, Magalie Ochs et Isabelle Régner, pour m'avoir suivie durant ces (presque) quatre années de thèse. J'ai eu beaucoup de chance d'avoir deux expertes dans leur domaine comme directrices. Je vous remercie pour tout ce que vous avez fait pour moi et pour m'avoir fait découvrir tant de choses passionnantes, comme le domaine des agents virtuels où j'ai pu programmer avec Unity pour la première fois et rendre « vivants » nos petits personnages, ou encore les joies des analyses statistiques et des interprétations qui vont avec. Merci encore !

J'aimerais adresser ma pleine reconnaissance et tous mes remerciements à Justine Cassell et Pascal Pansu pour avoir accepté de rapporter mon travail de thèse, ainsi qu'à Jean-Claude Martin et Agnès Blaye pour avoir accepté d'être examinateur et examinatrice. Je sais que c'est une chance unique de pouvoir discuter de mes travaux avec vous. Je remercie également mon comité de suivi de thèse, Domitile Lourdeaux et Johannes Ziegler, pour votre suivi et les pistes de réflexion que vous m'avez apporté pour mener à bien cette thèse.

Par ailleurs, je remercie Ampiric ainsi que ses membres, Mathilde Favier, Nicolas Mascret, Johannes Ziegler (encore une fois), Jonathan Mirault et tous les autres, sans qui cette thèse n'aurait jamais vu le jour. J'ai non seulement pu accéder plus facilement aux établissements grâce à eux, mais j'ai également pu bénéficier de leur aide précieuse pour le matériel d'expérimentation. Je remercie aussi l'ILCB pour son aide financière quant à la création des modèles 3D des agents virtuels. Je remercie également les établissements d'avoir accepté de participer aux études de cette thèse. Plus particulièrement, j'adresse toute ma reconnaissance et mes remerciements aux enseignants, aux référents, aux inspecteurs académiques ainsi qu'aux techniciens présents dans les collèges. Les études n'auraient pu être réalisées sans leur participation ! Merci beaucoup M. Brice, Mme Montreuil, Mme Chevance, Mme Espariat, Mme Durand, Mme Peres, M. Durand, M. Micoud, M. Fiaert, M. Bontoux, M. Boulangue, M. Saadaoui, Mme Grasso, M. Munos, Mme Archer, M. Drihem, M. Mazuyer, M. Deulofeu, Mme Bernard, M. Chekroun, Mme Laplace. Je m'excuse si j'ai oublié de citer certains d'entre vous. Cette thèse a été l'occasion de rencontrer d'excellents enseignants, tant humainement que professionnellement. Je remercie évidemment les élèves, et leur souhaite le meilleur pour la suite de leur scolarité, en espérant que ces jeunes enfants pourront choisir leur orientation professionnelle, libres de tout jugement concernant leur genre, leur origine culturelle ou encore leur origine sociale. J'espère de tout cœur qu'ils arriveront tous à atteindre leurs projets, qu'ils soient professionnels ou person-

nels.

Je remercie l'équipe ATYPICS du CRPN de m'avoir accueilli, ainsi que celle de TALEP du LIS. Il est vrai que j'ai passé beaucoup plus de temps au LIS, là où j'ai pu rencontrer de formidables personnes qui sont devenues pour la grande majorité mes amis. Merci beaucoup à mes amies qui partagent mon bureau, Alice et Marion ! Nos discussions pendant nos pauses vont beaucoup me manquer. Dans un monde idéal, j'aurais aimé qu'on continue de travailler au même endroit, dans le même bureau, avec la même petite (non-)plante ! Je remercie aussi tous les doctorants/post-docs de TALEP que j'adore et avec qui j'ai passé de bons moments, Elie, Eliot, Eunice, Franck, Ioana, Dhia, Afef, Emmanuelle, François (oui, je te compte dans les TALEPs). Vous avez tous contribué à rendre mes journées plus joyeuses. Je remercie également les permanents de l'équipe. Parmi eux, Alexis Nasr, Carlos Ramisch, et Benoit Favre ont été mes professeurs quand j'étais en licence ou en master. Je pense qu'ils ne se sont pas rendus compte, mais ils font partie des professeurs qui m'ont donné envie de faire de la recherche. Enfin, parmi le LIS, je ne pourrai jamais oublier les gestionnaires qui sauvent les doctorants tous les jours, et avec qui j'ai passé d'excellents moments, notamment pendant les heures de sport avec Jean-Pierre (que je remercie également). Merci Nadine, Sylvie, Patricia, et Nelly ! Et merci à toute l'équipe du sport !

Un grand merci à mes amies et à leurs soutiens pendant ces années de thèse : Clélia, Ilona, Zoë, Ally, Dilara. Votre soutien indéfectible m'a toujours permis d'avancer. J'ai eu énormément de chance de vous avoir rencontré, et j'ai encore plus de chance de vous compter parmi mes amies. Je remercie également les amis que j'ai rencontrés pendant mes études, en particulier Ibtissem, Loïc, et François-Xavier.

Également, je dois beaucoup à toute ma famille, en particulier à mes parents et à mon frère Yoann. Merci de m'avoir fait confiance toute ma vie et de m'avoir encouragée dans cette thèse. Merci d'être là pour moi. Il n'y aurait malheureusement pas assez de place pour citer tous les membres de ma famille. J'ai une pensée particulière pour mes grands-parents, je me suis souvent dit au cours de ma thèse « s'ils avaient été là, ils auraient été contents de ce que je fais ».

Enfin, je remercie Thomas, avec qui je partage ma vie depuis une douzaine d'années. Je te remercie pour tout ce que tu m'as apporté : ton aide précieuse, tes conseils, ton soutien, et surtout ton amour.

Ces travaux ont été soutenus par le pôle pilote pour la recherche en éducation Ampiric, une opération financée par l'État dans le cadre de l'action « Territoires d'innovation pédagogique » du Programme d'investissements d'avenir/France 2030, opéré par la Caisse des Dépôts.

Table des matières

Affidavit 2

Liste de publications et participations aux congrès et conférences 3

Résumé 6

Abstract 7

Remerciements 8

Table des matières 10

Table des figures 14

Liste des tableaux 17

Introduction générale 20

I. État de l'art 26

1. Introduction de la première partie 27

2. Les effets de « menace du stéréotype » sur les performances en mathématiques des filles et des femmes 29

2.1. Perceptions stéréotypées des compétences en mathématiques des filles et des garçons . . . 30
2.1.1. Les stéréotypes de genre : définition . . . 30
2.1.2. Des perceptions stéréotypées des compétences en mathématiques 31
2.1.3. Controverse autour de l'idéologie du don en mathématiques . . 32
2.2. Le phénomène de menace du stéréotype : définition et principaux mécanismes explicatifs . . . 34
2.3. Les effets de menace du stéréotype sur la performance des filles et des femmes en mathématiques . . . 37
2.3.1. Illustration de la menace du stéréotype en mathématiques . . . 37
2.3.2. Modérateurs des effets de menace du stéréotype . . . 40
2.3.3. Biais méthodologiques pouvant causer des problèmes de réplication . . . 46
2.3.4. Méthodes pour réduire les effets de menace du stéréotype . . . 50
2.4. Discussion . . . 61

2.5. Points clés du chapitre 63

3. Agents virtuels pédagogiques : modélisation du genre, du rôle de modèle de réussite, et du rapport virtuel 64
3.1. Introduction 65
3.2. Impact du genre des agents virtuels sur la perception et la performance des utilisateurs 68
3.2.1. Méthode de la revue systématique 68
3.2.2. Évaluations des agents virtuels 69
3.2.3. Persistance des stéréotypes de genre dans les interactions humain-agent 72
3.2.4. Effet du genre des agents virtuels pédagogiques sur les résultats scolaires 91
3.3. Agents virtuels pédagogiques en tant que modèles de réussite 102
3.4. Le rapport virtuel pour améliorer les performances académiques . . . 106
3.4.1. Fondement théorique du rapport 106
3.4.2. Modélisation du rapport virtuel 112
3.5. Vers un modèle d'agent virtuel pédagogique pour lutter contre la menace du stéréotype 118
3.6. Discussion 121
3.7. Points clés du chapitre 123

4. Problématique de recherche 124

II. Études empiriques 129

5. Introduction de la deuxième partie 130

6. Modélisation et implémentation des agents virtuels pédagogiques et de l'environnement virtuel d'apprentissage 133
6.1. Objectifs 134
6.2. Agents virtuels et environnement virtuel 134
6.2.1. Environnement virtuel de la salle de classe 135
6.2.2. Création des premiers modèles 3D des agents virtuels 138
6.2.3. Perception de l'âge des agents virtuels 139
6.2.4. Modèles 3D finaux des agents virtuels 140
6.3. Modélisation computationnel du comportement de rapport 142
6.3.1. Bibliothèque de comportements multimodaux de rapport . . . 143
6.3.2. Association des comportements non-verbaux et verbaux 157
6.4. Intégration des agents virtuels dans l'environnement virtuel d'apprentissage interactif 163
6.5. Création de trois types d'agents virtuels 171
6.6. Discussion 171
6.7. Points clés du chapitre 173

7. Études pilotes 1A et 1B : évaluation des agents virtuels animés 174
7.1. Objectifs . . . 175
7.2. Étude pilote 1A : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent . . . 177
7.2.1. Participants . . . 177
7.2.2. Procédure . . . 177
7.2.3. Mesures . . . 178
7.2.4. Résultats . . . 179
7.2.5. Discussion . . . 185
7.3. Étude pilote 1B : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent et de leurs explications de la leçon de mathématiques . . . 187
7.3.1. Participants . . . 187
7.3.2. Procédure . . . 187
7.3.3. Mesures . . . 187
7.3.4. Résultats . . . 187
7.3.5. Discussion . . . 195
7.4. Discussion . . . 196
7.5. Points clés du chapitre . . . 199

8. Étude pilote 2 : élaboration et évaluation des tests standardisés de mathématiques 200
8.1. Objectifs . . . 201
8.2. Matériels créés . . . 202
8.2.1. Classe de 6e . . . 202
8.2.2. Classe de 5e . . . 202
8.3. Participants . . . 203
8.4. Procédure . . . 203
8.5. Mesures et stratégies d'analyses descriptives . . . 204
8.6. Résultats . . . 205
8.6.1. Analyses descriptives globales . . . 205
8.6.2. Analyses descriptives chez les élèves de 6e . . . 206
8.6.3. Analyses descriptives chez les élèves de 5e . . . 210
8.7. Discussion . . . 214
8.8. Points clés du chapitre . . . 215

9. Étude de l'impact des agents virtuels pédagogiques sur la performance des élèves en mathématiques 216
9.1. Objectifs . . . 217
9.2. Méthode . . . 222
9.2.1. Participants . . . 222
9.2.2. Procédure . . . 222
9.2.3. Mesures . . . 224

9.3. Résultats . . . 231
9.3.1. Analyses préliminaires . . . 231
9.3.2. Analyses centrales . . . 241
9.4. Discussion . . . 247
9.5. Points clés du chapitre . . . 253

III. Discussion générale, limites, perspectives 255

10.Discussion générale 256
10.1. Principales contributions de la thèse . . . 257
10.2. Implications théoriques . . . 260
10.3. Implications pratiques . . . 262
10.4. Implications éthiques . . . 263
10.4.1. Apparence des agents virtuels pédagogiques . . . 264
10.4.2. Comportements verbaux et non-verbaux des agents virtuels pédagogiques . . . 265
10.4.3. Quelques réflexions personnelles . . . 267
10.5. Limites . . . 268
10.5.1. Limites des expérimentations . . . 268
10.5.2. Limites des agents virtuels pédagogiques . . . 269
10.6. Perspectives . . . 270

Bibliographie 272

ANNEXES 300
A. Liste des acronymes . . . 300
B. Autorisation du comité d'éthique . . . 301
C. Formulaire d'autorisation parentale . . . 302
D. Dialogues de l'agent avec rapport, avec les balises SSML . . . 303
E. Dialogues de l'agent sans rapport, avec les balises SSML . . . 308
F. Test standardisé de mathématiques 1 pour les 6e. . . . 312
G. Test standardisé de mathématiques 2 pour les 6e. . . . 316
H. Test standardisé de mathématiques 1 pour les 5e. . . . 322
I. Test standardisé de mathématiques 2 pour les 5e. . . . 328
J. Questionnaire d'évaluation des agents virtuels pédagogiques (études pilotes 1A et 1B du Chapitre 7.) . . . 333
K. TAI genre-mathématiques-lecture passé par les élèves dans l'étude principale (chapitre 9) . . . 336

Table des figures

3.1. Caractéristiques des agents expert, motivateur, et mentor (BAYLOR & KIM, 2005) . . . 103
3.2. Agents masculin et féminin de l'étude de SHIBAN et al. (2015) . . . 104
3.3. Éléments du rapport (E. MURPHY & RODRÍGUEZ-MANZANARES, 2012). 107
3.4. Importance des trois composantes du rapport à travers le temps (TICKLE-DEGNEN & ROSENTHAL, 1990). . . . 109
3.5. Stratégies de communication et fonctions sociales pour construire et maintenir le rapport (R. ZHAO et al., 2014). . . . 115
3.6. Stratégies de communication et fonctions sociales baissant le niveau de rapport (R. ZHAO et al., 2014). . . . 115
3.7. Exemples de dialogues entre 2 élèves ne se connaissant pas (à gauche) et entre 2 élèves amis (à droite) provenant du corpus de R. ZHAO et al. (2014).116

6.1. Illustrations de tous les éléments de décoration en 3D présents dans la salle de classe virtuelle. . . . 136
6.2. Salle de classe virtuelle. . . . 137
6.3. Différents point de vue de la scène 3D de la salle de classe virtuelle sous Unity 3D. . . . 137
6.4. Premières versions de l'agent féminin. . . . 138
6.5. Photos libres de droits d'adolescents réels à gauche, dont le visage a inspiré les agents à droite. . . . 139
6.6. Images fixes des agents évalués par une classe de 6e. . . . 140
6.7. Modèles 3D des agents virtuels finaux. . . . 141
6.8. Modèles 3D des agents virtuels finaux dans l'environnement virtuel, assis sur une chaise face à la caméra. . . . 142
6.9. Illustrations des différents sourires. . . . 145
6.10. Illustrations des différentes expressions des sourcils. . . . 146
6.11. L'agent demande à l'élève : « C'est quoi ton nom? ». L'élève utilise le clavier virtuel sur tablette, et le clavier sur ordinateur. . . . 149
6.12. Différents gestes portés vers soi-même. . . . 149
6.13. L'agent demande à l'élève : « Ça va? Tu t'en sors bien en maths? ». L'élève doit cliquer sur une des quatre réponses proposées (« oui! », « Des fois oui, des fois non. », « Bof. », « Non! »). L'agent s'adapte alors à la réponse de l'élève. . . . 151
6.14. Exemple de dialogues où l'agent fait référence à une expérience partagée par l'élève. . . . 151
6.15. Illustrations de différents gestes amples et ouverts. . . . 152

6.16. Tous les gestes déictiques. 154
6.17. Animation du geste iconique de l'addition. 155
6.18. Animation du geste iconique de la soustraction. 155
6.19. Animation du geste iconique de la division. 155
6.20. Animation du geste iconique de la multiplication. 156
6.21. Gestes iconiques des droites sécantes (a), des droites parallèles (b), et des droites perpendiculaires (c). 156
6.22. Bras légèrement levé avec l'index tendu pour souligner un point du cours présenté verbalement. 157
6.23. Présentation de l'agent virtuel avec des comportements de rapport virtuel. 158
6.24. Suite de la présentation de l'agent virtuel avec des comportements de rapport virtuel. 159
6.25. Modules des agents virtuels. 164
6.26. Tous les *visemes* créés pour l'animation labiale avec le module SALSA. 165
6.27. Interface d'Unity pour créer une animation, ici le geste iconique pour expliquer la division. Le Panel A montre l'animation représentée par des points, et le Panel B montre la même animation représentée par des courbes. 166
6.28. Combinaison d'un geste ouvert et d'un geste porté vers soi-même, en utilisant des masques. 167
6.29. Masques créés et utilisés pour animer les agents virtuels. 168
6.30. Interface sous Unity montrant une interaction avec l'agent virtuel, et les packages pour animer le personnage (*Salsa* et *Character*) et pour gérer l'interface utilisateur (*UI*). 170

7.1. Étude pilote 1B. Score moyen du sentiment de rapport virtuel en fonction du sexe de l'élève (fille ou garçon) et du genre de l'agent (féminin ou masculin). 192
7.2. Étude pilote 1B. Score moyen du niveau d'anthropomorphisme en fonction du comportement de l'agent (avec ou sans rapport) et du genre de l'agent (féminin ou masculin). 193

8.1. Étude pilote 2. Score moyen des performances des garçons et des filles de 6e pour chaque exercice du test 1. 206
8.2. Étude pilote 2. Exercice (jugé difficile par les enseignants) rajouté au test 1 de 6e. 207
8.3. Étude pilote 2. Score moyen des performances des garçons et des filles de 6e pour chaque exercice du test 2. 208
8.4. Étude pilote 2. Exercice supprimé au test 2 de 6e à cause de sa facilité. . 209
8.5. Étude pilote 2. Exercices (jugés difficiles par les enseignants) rajoutés au test 2 de 6e. 210
8.6. Étude pilote 2. Score moyen des performances des garçons et des filles de 5e pour chaque exercice du test 1. 211
8.7. Étude pilote 2. Exercices supprimés au test 1 de 5e à cause de leur facilité. 212

8.8. Étude pilote 2. Score moyen des performances des garçons et des filles de 5e pour chaque exercice du test 2. 213
8.9. Étude pilote 2. Exercices supprimés au test 1 de 5e à cause de leur facilité.214

9.1. Schéma explicatif des tâches de l'étude principale. 221
9.2. Exemples du TAI genre-mathématiques-lecture où l'on doit classer les mots avec les touches E et I du clavier (a) ou avec les flèches sur l'écran tactile de la tablette (b). 226
9.3. Exemples du TAI genre-mathématiques-lecture de la condition congruente (a) et incongruente (b) avec le stéréotype. 227
9.4. Exemples du TAI genre-mathématiques-lecture donnés au tableau de la salle, lorsqu'il n'y a que deux catégories (panel A) et lorsque les quatre catégories sont présentes (panel B). 229
9.5. Étude principale. Moyennes trimestrielles en mathématiques des filles et des garçons selon leurs classes, notées de manière classique (note sur 20). 234
9.6. Étude principale. Effectif des filles et des garçons dans les trois groupes de niveaux de la notation en couleur. 235
9.7. Étude principale. Score moyen du sentiment de réussite en mathématiques chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves, au temps 1 (panel A) et au temps 2 (panel B). 237
9.8. Étude principale. Score moyen de l'importance accordée aux mathématiques chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves, au temps 1 (panel A) et au temps 2 (panel B). 238
9.9. Étude principale. Score moyen au TAI genre-mathématiques-lecture au temps 2 chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves. 239
9.10. Étude principale. Score moyen des performances globales chez les filles et les garçons au temps 1, en fonction de la voix de la leçon. 242
9.11. Étude principale. Score moyen des performances chez les filles et les garçons au temps 1 sur les exercices faciles (panel A) et difficiles (panel B), en fonction de la voix de la leçon. 243
9.12. Étude principale. Score moyen des performances globales des élèves au temps 2, en fonction du type de l'agent et de son genre. 245

1. Capture d'écran des instructions globales pour effectuer le TAI sur ordinateur. 337
2. Captures d'écran des instructions de chaque block pour effectuer le TAI sur ordinateur. 338
3. Capture d'écran des instructions globales pour effectuer le TAI sur tablette.339
4. Captures d'écran des instructions de chaque block pour effectuer le TAI sur tablette. 340

Liste des tableaux

2.1. Participants, inductions expérimentales, et résultats obtenus dans chacune des trois études de GANLEY et al. (2013). 48

3.1. Synthèse d'articles sur les études perceptives d'agents virtuels en fonction de leur genre 73
3.2. Synthèse d'articles sur l'impact des agents virtuels genrés dans le cadre d'une tâche d'apprentissage 92

5.1. Résumé de chacune de nos études. Les noms des collèges ont été anonymisés. 132

6.1. Expressions faciales avec leur amplitude (de 0 à 100), leurs *Blendshapes*, et l'amplitude de chaque *Blendshape*. 147
6.2. Description des étiquettes d'animations utilisées dans les Figures 6.23 et 6.24. 160

7.1. Étude pilote 1A. Score moyen (erreur standard) pour l'ensemble des variables en fonction du genre de l'élève, du genre de l'agent et du comportement de l'agent avec ou sans rapport virtuel. 180
7.2. Étude pilote 1A. Corrélations bivariées entre l'ensemble des variables dans la condition sans rapport virtuel (panel A) et dans la condition avec rapport virtuel (panel B). 181
7.3. Étude pilote 1A. Score moyen et erreur standard pour chaque variable dépendante, en fonction du comportement de l'agent (avec ou sans rapport). 183
7.4. Étude pilote 1A. Analyses de régressions multiples réalisées séparément dans la condition sans rapport virtuel *versus* avec rapport virtuel. . . . 184
7.5. Étude pilote 1B. Score moyen (erreur standard) pour l'ensemble des variables en fonction du genre de l'élève, du genre de l'agent et du comportement de l'agent avec ou sans rapport virtuel. 188
7.6. Étude pilote 1B. Corrélations bivariées entre l'ensemble des variables dans la condition sans rapport virtuel (panel A) et dans la condition avec rapport virtuel (panel B). 189
7.7. Étude pilote 1B. Score moyen et erreur standard pour chaque variable dépendante, en fonction du genre de l'agent (féminin ou masculin). . . 191
7.8. Étude pilote 1B. Analyses de régressions multiples réalisées séparément dans la condition sans rapport virtuel versus avec rapport virtuel. . . . 194

8.1. Étude pilote 2. Nombre de participants avec leur moyenne trimestrielle en mathématiques (et erreur standard) par conditions expérimentales. 204
8.2. Étude pilote 2. Moyennes (et erreurs standards) trimestrielle en mathématiques et obtenues lors des tests standardisés de mathématiques, pour les filles et les garçons. 205
8.3. Étude pilote 2. Moyenne et erreur standard pour chaque exercice du test 1 des 6e. 206
8.4. Étude pilote 2. Moyenne et erreur standard pour chaque exercice du test 2 des 6e. 208
8.5. Étude pilote 2. Moyenne et erreur standard pour chaque exercice du test 1 des 5e. 211
8.6. Étude pilote 2. Moyenne et erreur standard pour chaque exercice du test 2 des 5e. 213

9.1. Étude principale. Nombre de participants par conditions expérimentales au temps 1, après avoir appliqué les critères d'exclusions. 222
9.2. Étude principale. Nombre de participants par conditions expérimentales présents à la fois aus temps 1 et 2, après avoir appliqué les critères d'exclusions. 223
9.3. Étude principale. Scores moyens, erreurs standards, et corrélations bivariées entre l'ensemble des mesures. 232
9.4. Étude principale. Nombre de participants selon le type de notation, aux temps 1 et 2. 233
9.5. Étude principale. Scores moyens et erreurs standards pour chaque score de performance, au temps 1 et au temps 2, selon le type de notation en classe (classique ou couleur). 236
9.6. Étude principale. Nombre de participants inclus dans les analyses au temps 2, par conditions expérimentales 244

1. Items utilisés pour la mesure de la compétence en mathématiques perçue de l'agent virtuel pédagogique. 333
2. Items utilisés pour la mesure du niveau d'anthropomorphisme perçu de l'agent virtuel pédagogique, adaptés du questionnaire du Godspeed évaluant l'anthropomorphisme et celle évaluant l'animation de BARTNECK et al. (2009). 333
3. Items utilisés pour la mesure de la sympathie perçue de l'agent virtuel pédagogique, adaptés du questionnaire du Godspeed de BARTNECK et al. (2009). 334
4. Items utilisés pour la mesure de l'identification à l'agent virtuel pédagogique, adaptés de O'BRIEN et al. (2017). 334
5. Items utilisés pour la mesure de l'identification à la réussite en mathématiques de l'agent virtuel pédagogique. 334

6. Items utilisés pour la mesure du sentiment de rapport virtuel perçu lors de l'interaction avec l'agent virtuel pédagogique, adaptés du questionnaire du Virtual Scale Intimacy (VIS) de POTDEVIN (2020) et du questionnaire de la perception de l'interaction de CUPERMAN et ICKES (2009). 335

Introduction générale

"*Il existait une masse immense de déclarations masculines tendant à démontrer qu'on ne pouvait rien attendre, intellectuellement, d'une femme. Même si son père ne lui faisait pas lecture de ces déclarations, toute fille pouvait les lire, et cette lecture, même au XIXe siècle, devait diminuer sa vitalité et affecter profondément son œuvre. Quoi qu'il arrivât, cette affirmation existait –vous ne pouvez faire ceci, vous êtes incapable de faire cela– et il fallait protester contre elle, il fallait la surmonter.*" Virginia Woolf (Une chambre à soi).

Bien que la citation ci-dessus date du XIXe siècle, un grand nombre d'études en sciences sociales montre que cette observation est toujours d'actualité. Les stéréotypes de genre confinent toujours les femmes et les hommes dans des rôles, et ils continuent d'alimenter et d'être alimentés par les observations sur les écarts de performances en mathématiques en faveur des garçons et des hommes. Ces écarts existent bel et bien dans la majorité des pays de l'OCDE (Organisation de Coopération et de Développement Économiques) et d'autres pays non-membres (BERNIGOLE et al., 2023). En France, la dernière évaluation de PISA en 2022 (*Programme for International Student Assessment*, ou programme international pour le suivi des acquis des élèves de 15 ans), mais également des évaluations précédentes (FUMEL et al., 2016), montrent un écart de performance significatif en mathématiques en faveur des garçons (BERNIGOLE et al., 2023). En France en 2022, les garçons ont obtenu un score de 479 et les filles de 469. À noter qu'il n'y avait pas de différence significative de performance entre les filles et les garçons dans seulement 10 pays comme la Belgique, la Grèce, ou encore la Turquie. Les filles ont obtenu de meilleurs résultats en mathématiques que les garçons uniquement en Finlande. L'écart se réduit en France, mais s'explique par une baisse de performance des garçons, et non pas par une amélioration des résultats des filles. De plus, en France, les filles sont beaucoup plus anxieuses que les garçons en mathématiques. La France est, avec l'Allemagne et la Norvège, le pays où cette différence d'anxiété entre les filles et les garçons est la plus grande : 55.6% des garçons déclarent avoir peur d'obtenir des mauvaises notes en mathématiques, contre 71.9% des filles (BERNIGOLE et al., 2023).

Des écarts comparables de performance en faveur des garçons ont été constatés dans d'autres évaluations, comme celles du TIMMS Advanced (*Trends in International Mathematics and Science Study Advanced*, ou étude de l'évolution des acquis internationaux en mathématiques et en sciences avancée) avec les élèves français de Terminale S (SALLES & LE CAM, 2016). **Parmi les 9 pays ayant participé au TIMSS Advanced 2015, la France est le pays où l'écart de performance entre filles et garçons**

est le plus grand, en proportion d'écart-type : les garçons ont obtenu un score de 475, contre 449 chez les filles, correspondant à 36.6% de l'écart-type des scores. Chez les premières années en CPGE (Classes Préparatoires aux Grandes Écoles) scientifiques, on aurait pu penser que l'écart de performance entre les filles et les garçons se réduirait, puisqu'il faut de bons résultats scolaires pour y accéder. Cependant, ce n'est pas ce qui est observé : le taux moyen de réussite est de 65% pour les filles contre 70% pour les garçons en mathématiques. Une étude longitudinale de J.-P. FISCHER et THIERRY (2022) des résultats scolaires de 18000 enfants à partir de la maternelle en France durant 20 ans montre que les écarts de performances en mathématiques en faveur des garçons émergent dès 6-7 ans, au CP, puis se poursuivent tout au long de la scolarité. Jusqu'à l'âge de 4-5 ans, il y a un léger avantage non-significatif pour les filles. Cette différenciation de résultats débute plus tôt dans les milieux sociaux favorisés. Au CP, il n'y a ainsi aucune différence de résultats en mathématiques entre les filles et les garçons dans les milieux très défavorisés, alors qu'on observe un écart de 10% en faveur des garçons dans les milieux très favorisés. ANDREU et al. (2022) ont observé, en 2021, que les garçons obtenaient de meilleurs résultats par rapport aux filles en mathématiques lors des évaluations d'entrée en 6e. Le constat est le même dans le dernier rapport de la Direction de l'Evaluation, de la Prospective et de la Performance (DEPP) en 2022. De plus, en mathématiques, 35.1% des filles se situaient dans les groupes de bas niveaux contre 29.9% des garçons. Entre 2017 et 2022, la proportion d'élèves dans les groupes de bas niveaux n'a pas évolué pour les garçons alors que la part des filles a augmenté de 3.6 points. Ce sont parmi les groupes de hauts niveaux que la proportion de garçons a augmenté de 5.2 points, contre 1.0 point pour les filles (ANDREU et al., 2023). Ces écarts s'observent en particulier sur les tests standardisés tels que les évaluations nationales, mais ne se retrouvent pas toujours en cours où les filles ont généralement des résultats similaires aux garçons (HYDE & LINN, 2006 ; KIMBALL, 1989).

Ce sont plus particulièrement des études en psychologie cognitive et sociale qui s'intéressent aux explications quant aux écarts de performance en mathématiques en faveur des garçons et des hommes, notamment à travers les travaux portant sur la **menace du stéréotype**. Ce phénomène, mis en évidence par C. M. STEELE et ARONSON (1995) aux États-Unis, illustre l'impact des stéréotypes sur les performances des groupes ciblés lors de test de compétences. En ayant peur de confirmer un stéréotype négatif ciblant un groupe social auquel l'individu appartient, ou bien en ayant peur d'être lui-même négativement stéréotypé, l'individu en question peut produire des contre-performances lors d'un test à fort enjeu. C'est en 1999 que des chercheurs ont observé l'impact du stéréotype portant sur les compétences en mathématiques des femmes sur leurs performances dans ce domaine. Le stéréotype de genre en sciences, toujours largement diffusé dans nos sociétés, véhicule l'idée que les filles et les femmes seraient naturellement moins douées que les garçons et les hommes en mathématiques. Ce stéréotype fourni donc un argument qui est très fréquemment avancé pour expliquer la faible représentation des filles et des femmes dans les domaines liés aux mathématiques, ainsi que leurs performances parfois moins élevées

dans les évaluations standardisées, en occultant d'autres explications potentielles. Pourtant, SPENCER et al. (1999) ont montré que l'on peut réduire les effets de ce stéréotype, grâce à l'ajout d'une simple instruction avant le test qui mentionne que les hommes et les femmes obtiennent les mêmes résultats au test de mathématiques à venir. En rendant le stéréotype non pertinent sur le test en question, cette consigne a permis aux femmes d'obtenir d'aussi bons résultats que les hommes à un test de mathématiques très complexe. Pour la première fois, les explications biologisantes au cœur du stéréotype de genre ont été déstabilisées. *La menace du stéréotype peut en effet expliquer, en partie, pourquoi les filles et les femmes ont de moins bons résultats lors de tests standardisés de mathématiques complexes, comparativement aux garçons et aux hommes.* Dans cette thèse, nous nous intéressons spécifiquement à la menace du stéréotype envers les filles en mathématiques.

Encourager l'égalité filles-garçons dans l'éducation et en particulier en mathématiques passe donc notamment par des outils pour réduire les effets de menace du stéréotype. Ce phénomène agit dès l'école, même chez les jeunes enfants au CP (RÉGNER et al., 2014). Il est important de s'en préoccuper pour permettre aux filles de prendre conscience qu'elles peuvent réussir en mathématiques grâce à l'amélioration de leurs performances, qu'elles prennent confiance en elles dans ce domaine et éviter un sentiment de non-appartenance (voire de désidentification aux mathématiques), afin qu'elles puissent plus facilement s'orienter vers des études et des métiers en Sciences, Technologie, Ingénierie et Mathématiques (STIM). Cette préoccupation constitue un véritable enjeu sociétal, où l'école a pour responsabilité de promouvoir et garantir l'égalité des chances entre les filles et les garçons, comme indiqué par les articles L121-1 et L312-17-1 du Code de l'éducation qui soulignent que l'école doit œuvrer à encourager la mixité et l'égalité entre les femmes et les hommes à tous les niveaux, notamment en ce qui concerne l'orientation scolaire.

La recherche en psychologie sociale a montré que les différences de performance en mathématiques entre filles et garçons ne peuvent être simplement attribuées à des différences innées, mais sont le résultat de multiples influences sociales et culturelles, parmi lesquelles les stéréotypes de genre jouent un rôle majeur. Les études montrent qu'il est ainsi possible de diminuer voire d'éliminer les disparités de résultats en mathématiques entre filles et garçons, en intervenant avant que les élèves ne passent un test de mathématiques. Plusieurs méthodes se sont révélées efficaces pour lutter contre la menace du stéréotype. La présentation de **modèles de réussite** incarnés par des filles ou des femmes compétentes en mathématiques fait partie des méthodes efficaces pour réduire les effets de menace du stéréotype (MARX & KO, 2012). En cas d'anxiété, comme celle engendrée par une tâche jugée trop difficile, une stratégie consiste à rechercher une affiliation en s'identifiant à quelqu'un d'autre. Cette personne devient ainsi un modèle de réussite, recherché pour ses attributs tels que le statut et les compétences. Les apprenants ont tendance à s'attacher à des individus légèrement plus âgés et plus élevés en statut, tout en partageant des similitudes comme le genre ou l'origine culturelle. L'objectif de l'apprenant est alors de s'inspirer de cette

personne et de lui ressembler. Plusieurs études ont montré qu'une seule exposition à un modèle de réussite en sciences qui explique sa réussite par le travail et les efforts permettait de réduire les effets de menace du stéréotype (BAGÈS & MARTINOT, 2011 ; BAGÈS et al., 2016). Cependant, une seule exposition ne garantit pas de meilleurs résultats en mathématiques sur le long terme. Il est cependant impossible de faire intervenir des modèles de réussite dans les classes, tout le temps, juste avant de passer un test de mathématiques. Une méthode, encore non explorée et qui constitue l'objectif de la thèse, pourrait consister à utiliser des personnages virtuels pour incarner des modèles de réussite auprès des jeunes filles, à travers des applications pédagogiques.

Les applications éducatives sont des plateformes numériques pouvant fournir des exercices comme des cours dans un domaine académique. De plus en plus utilisé en classe comme à la maison, on dénombre 455 000 applications éducatives dans l'App Store d'Apple et sur Google Play en 2022. La même année, le domaine de l'*EdTech* a généré $7 billions de revenu, avec une croissance accélérée pendant la pandémie de la Covid-19. Au-delà de cette crise sanitaire, le taux de croissance annuel moyen de l'*EdTech* devrait augmenter de 8.9% jusqu'en 2030 (WYLIE, 2024). Pourtant, nombre de ces applications ne sont pas évaluées (HIRSH-PASEK et al., 2015). Il n'est pas déraisonnable de faire l'hypothèse que les environnements virtuels d'apprentissage de ces applications n'échappent pas à la menace du stéréotype. Comme énoncé plus haut, les personnages virtuels parfois présents dans ces applications, appelés **agents virtuels pédagogiques**, pourraient incarner des modèles de réussite de sorte à réduire les effets de menace du stéréotype. Les agents virtuels pédagogiques sont une forme spécifique d'*agent virtuel incarné*. Un agent virtuel incarné, également connu sous le nom d'Agent Conversationnel Animé (ACA), est une interface où le système est représenté par un personnage ou une entité autonome, permettant la transmission d'informations aux utilisateurs par diverses modalités (voix, gestes, expressions faciales) (CASSELL, 2001 ; LUGRIN et al., 2022). Plus précisément, les agents virtuels *pédagogiques* sont des ACA utilisés dans des environnements numériques pour améliorer l'apprentissage en milieu éducatif (MOHTADI et al., 2014). L'objectif principal de ces agents virtuels pédagogiques est d'enrichir l'expérience des utilisateurs dans un contexte éducatif, en les assistant et en les orientant. Ces agents virtuels sont connus pour améliorer l'apprentissage des utilisateurs ainsi que leur motivation, en particulier lorsqu'ils incarnent le rôle pédagogique de *mentor* où ils peuvent montrer leurs connaissances tout en encourageant les élèves (BAYLOR & KIM, 2005 ; Y. KIM & BAYLOR, 2016). Cependant, à ce jour et à notre connaissance, **très peu d'étude a créé des agents virtuels pédagogiques pour qu'ils incarnent des modèles de réussite afin d'encourager les filles en mathématiques.** Ces quelques études n'ont cependant pas montré la capacité des agents à incarner des modèles de réussite ; elles ont plutôt postulé qu'ils étaient perçus comme tels. De tels agents virtuels pédagogiques pourraient répondre à deux des objectifs du ministère de l'Éducation nationale et de la Jeunesse énoncés en 2023 : celui de promouvoir l'égalité filles-garçons en mathématiques, et celui de "fournir aux professeurs une offre claire mêlant outils et ressources numériques pour mettre davantage le numérique au service de la réussite des élèves"

(stratégies numériques pour l'éducation 2023-2027).

L'objectif de notre travail de thèse est de concevoir des agents virtuels pédagogiques qui incarnent des modèles de réussite afin de réduire les effets de menace du stéréotype en mathématiques dans les classes. Nous avons élaboré des modèles computationnels de comportements humains variés pour évaluer l'impact de ces agents sur les effets de menace du stéréotype dans des conditions réelles d'utilisation en classe, en dehors du cadre du laboratoire. Nous visons à répondre à un ensemble de questions de recherche, à savoir :

- *Comment créer des agents virtuels pédagogiques autonomes pour qu'ils incarnent un modèle de réussite virtuel ?*
- *Quelles sont les caractéristiques comportementales d'un agent virtuel permettant une identification de la part des enfants ?*
- *Comment élaborer un modèle computationnel des comportements d'un agent pédagogique pour simuler un modèle de réussite auquel les enfants peuvent s'identifier ?*
- *Quel est l'impact de ces agents virtuels pédagogiques sur la performance en mathématiques des élèves, en particulier des filles ?*

Notre recherche s'inscrit dans le cadre du pôle pilote Ampiric (Aix-Marseille – Pôle d'Innovation, de Recherche, d'enseIgnement pour l'éduCation), dont l'objectif est d'améliorer durablement l'apprentissage des savoirs fondamentaux des élèves tout au long de leur parcours scolaire, en particulier ceux rencontrant des difficultés. Nos travaux visent en effet à améliorer les performances des élèves dans l'apprentissage des savoirs fondamentaux en mathématiques en développant, testant, et diffusant une approche pédagogique innovante fondée sur les technologies numériques : des agents virtuels pédagogiques pour réduire les effets de menace du stéréotype. Plus particulièrement, cette thèse répond à deux objectifs stratégiques d'Ampiric :

- **Améliorer la compréhension des processus d'enseignement-apprentissage des savoirs fondamentaux à travers des recherches fondamentales.** Les résultats attendus de cette thèse permettront de mieux identifier la façon dont les outils numériques peuvent être utilisés pour lutter contre la menace du stéréotype lors des apprentissages et des évaluations, et réduire les inégalités scolaires entre les élèves.
- **Coproduire des outils et des applications au service des apprentissages fondamentaux.** De cette thèse résultera le développement d'un prototype d'application avec un agent virtuel pédagogique pour améliorer l'apprentissage des mathématiques.

Ce manuscrit est organisé en trois parties. La première partie est consacrée à l'état de l'art sur la menace du stéréotype chez les filles et les femmes en mathématiques, et sur les agents virtuels pédagogiques. Plus particulièrement, nous montrons d'une part

que les stéréotypes de genre s'invitent également dans les interactions humain-agent, et d'autre part comment modéliser ces interactions afin qu'elles soient plus positives et harmonieuses. Le dernier chapitre de cette partie met en lumière l'interdisciplinarité de notre approche en mettant en relation les deux domaines de recherche pour aboutir à la création collaborative des agents virtuels pédagogiques. En effet, en intégrant les travaux de la psychologie sociale et de l'informatique, nous avons pu développer une approche de modélisation unique qui aurait été impossible à réaliser en se limitant à un seul champ disciplinaire.

La deuxième partie porte sur nos contributions empiriques. Nous présentons tout d'abord les agents virtuels pédagogiques que nous avons développés, créés spécifiquement pour réduire les effets de menace du stéréotype chez les filles en mathématiques. Nous détaillons le modèle computationnel qui a permis de déterminer les comportements multimodaux des agents virtuels. Ces agents ont été évalués auprès d'enfants dans deux études pilotes perceptives, nous permettant d'explorer un certain nombre de questions de recherche sur les caractéristiques et capacités nécessaires pour que les agents virtuels incarnent de meilleurs modèles de réussite. Enfin, nous exposons notre étude finale sur l'impact de ces agents sur la performance en mathématiques des élèves de 6e et de 5e.

Pour conclure, la troisième partie de la thèse discute des résultats et des limites de nos études, et propose des perspectives pour la recherche sur les agents virtuels pédagogiques et sur la menace du stéréotype.

Première partie

État de l’art

1. Introduction de la première partie

Dans l'objectif de créer des agents virtuels pédagogiques pour améliorer l'apprentissage des mathématiques, la littérature sur la menace du stéréotype est incontournable, car elle apporte des éléments de compréhension sur les effets des stéréotypes de genre défavorables aux filles et aux femmes dans les domaines des STIMs sur leurs performances, mais aussi des solutions potentielles pour réduire ces effets négatifs et ainsi aller vers plus d'égalité entre les filles et les garçons en sciences. La prise en compte des effets de ces stéréotypes de genre doit être faite dans les outils du numérique éducatif, hors ou dans la salle de classe. En effet, ces applications sont de plus en plus utilisées : par exemple, les cours d'introduction à la programmation sont très souvent aidés par le site https://studio.code.org/ pour que les élèves puissent directement tester et apprendre. Ainsi, pour réduire les effets des stéréotypes de genre dans un contexte numérique, la littérature sur la menace du stéréotype ne peut pas faire l'économie de la littérature sur les agents virtuels pédagogiques, car c'est l'informatique qui apporte l'aide aux élèves via ces agents virtuels. Aujourd'hui, le numérique éducatif est au cœur des apprentissages des enfants. Ces applications n'utilisent pas forcément d'agents virtuels pédagogiques, alors que la recherche en informatique montre l'efficacité de ces agents selon leurs rôles.

Cependant, un nombre considérable d'assistants vocaux et d'agents virtuels, pédagogiques ou non (e.g., hors contexte académique), a toutefois parfois occulté voire reproduit les stéréotypes de genre. Les concepteurs utilisent généralement des stéréotypes de genre pour la représentation physique et les réponses verbales des agents virtuels conversationnels (Brahnam & Weaver, 2015; Loideain & Adams, 2020; Strengers & Kennedy, 2021; West et al., 2019). Le rapport de l'UNESCO a par exemple révélé que les agents et assistants virtuels féminins font face à des préjugés sexistes et réagissent en adoptant des phrases passives, tolérantes et serviables, comme l'indique le titre de leur article (West et al., 2019). Il est donc nécessaire pour la recherche sur les agents virtuels pédagogiques de se nourrir des travaux sur la menace du stéréotype, et vice-versa, il est nécessaire pour la littérature sur la menace du stéréotype de connaître ce qu'il est possible de créer avec les agents virtuels pédagogiques. En effet, les travaux sur la menace du stéréotype doivent prendre en compte les agents virtuels pédagogiques dans l'amélioration de l'apprentissage des mathématiques pour être également efficaces dans un contexte numérique, et ainsi trouver de nouvelles interventions pour lutter contre la menace du stéréotype. De plus, la plupart des méthodes peuvent être implémentées avec un agent virtuel pédagogique (e.g., phrase de falsification énoncée par l'agent, éviter un habillage genré des exercices), où l'agent pourrait devenir un guide bienveillant pour les élèves en incarnant un modèle

de réussite en mathématiques. Quelles seraient les caractéristiques de l'agent pour atteindre cet objectif? La littérature sur la menace du stéréotype peut nous éclairer sur la question. C'est pour cela que nous commençons tout d'abord par un état de l'art sur la menace du stéréotype chez les filles et les femmes en mathématiques, pour comprendre comment les stéréotypes impactent les élèves (Chapitre 2). Comme nous l'avons expliqué, au-delà de comprendre les mécanismes cognitifs, la recherche a étudié les méthodes efficaces pour en réduire ses effets. Parmi elles, l'exposition à un modèle de réussite en mathématiques permet de faire le lien avec la littérature sur les agents virtuels pédagogiques et de répondre à la question en étudiant quels profils de modèles de réussite sont efficaces pour réduire les effets de menace du stéréotype. Les études dans le domaine des agents virtuels pédagogiques nous permettent d'en savoir plus sur la perception stéréotypée des utilisateurs envers les agents virtuels, l'impact du genre des agents virtuels sur les performances académiques, et comment construire leurs comportements pour favoriser une meilleure interaction avec l'utilisateur (Chapitre 3). Ces interactions positives, connues sous le nom de *rapport*, ont été étudiées dans les relations interhumaines comme dans les interactions agent-humain. Plus particulièrement, quelques études ont montré que le rapport *virtuel* entre un agent virtuel pédagogique et un apprenant permettait d'améliorer leurs performances. Ainsi, comment lier ces deux littératures dans un contexte interdisciplinaire pour aider l'apprentissage des mathématiques sans stéréotypes de genre?

L'enjeu de cette thèse est donc de pouvoir combiner les revues des littératures présentées dans cette première partie, afin de proposer des agents virtuels pédagogiques incarnant des modèles de réussite interactifs en créant un modèle computationnel de rapport virtuel. Ces agents seront alors évalués dans un premier temps sur des mesures perceptives telles que leurs compétences en mathématiques ou encore le degré d'identification des élèves envers ces agents. Ce sera également l'occasion de tester l'influence du rapport virtuel sur ces mesures. Enfin, nous testerons les agents directement dans les classes, afin d'examiner leurs influences sur les performances en mathématiques des élèves, et plus particulièrement des filles. L'hypothèse principale de cette thèse est qu'un agent virtuel pédagogique incarnant un modèle de réussite en mathématiques et exprimant des comportements de rapport virtuel permettrait de réduire les effets de menace du stéréotype. Le dernier chapitre de cette partie (Chapitre 4) éclaire notre processus de recherche ainsi que la progression de notre réflexion qui nous a conduites à formuler l'hypothèse précédemment énoncée.

2. Les effets de « menace du stéréotype » sur les performances en mathématiques des filles et des femmes

Sommaire

2.1. Perceptions stéréotypées des compétences en mathématiques des filles et des garçons . . . 30
2.1.1. Les stéréotypes de genre : définition . . . 30
2.1.2. Des perceptions stéréotypées des compétences en mathématiques 31
2.1.3. Controverse autour de l'idéologie du don en mathématiques . . 32
2.2. Le phénomène de menace du stéréotype : définition et principaux mécanismes explicatifs . . . 34
2.3. Les effets de menace du stéréotype sur la performance des filles et des femmes en mathématiques . . . 37
2.3.1. Illustration de la menace du stéréotype en mathématiques . . . 37
2.3.2. Modérateurs des effets de menace du stéréotype . . . 40
2.3.2.1. Identification aux mathématiques . . . 40
2.3.2.2. Identification à son groupe de genre . . . 41
2.3.2.3. Associations implicites stéréotypées . . . 41
2.3.2.4. Hostilité perçue de l'environnement . . . 43
2.3.2.5. Capacité en mémoire de travail . . . 44
2.3.2.6. Vulnérabilité à la pression . . . 45
2.3.2.7. Difficulté des exercices . . . 45
2.3.3. Biais méthodologiques pouvant causer des problèmes de réplication . . . 46
2.3.4. Méthodes pour réduire les effets de menace du stéréotype . . . 50
2.3.4.1. Réduction de la perception de menace . . . 51
2.3.4.2. Stratégies de *coping* . . . 53
2.3.4.3. Création d'un environnement sain pour protéger l'identité de genre : les modèles de réussite . . . 56
2.4. Discussion . . . 61
2.5. Points clés du chapitre . . . 63

Dans ce chapitre, nous proposons un état de l'art des études en cognition sociale qui montrent que les stéréotypes de genre peuvent contribuer à expliquer les différences de performance en mathématiques en faveur des garçons. Après avoir défini la notion de stéréotypes et plus spécifiquement celle de stéréotypes de genre, nous examinons comment ces stéréotypes peuvent influencer la perception de la réussite en mathématiques selon le genre des élèves. Ensuite, nous passons en revue les études empiriques qui démontrent l'impact négatif de ces stéréotypes sur les performances des filles et des femmes en mathématiques, un phénomène connu sous le nom de « menace du stéréotype ». Enfin, nous présentons les méthodes qui permettent de réduire les effets de menace du stéréotype en mathématiques, en insistant plus particulièrement sur la stratégie qui consiste à mettre en avant des modèles de réussite en sciences, puisqu'ils sont au cœur de cette thèse.

2.1. Perceptions stéréotypées des compétences en mathématiques des filles et des garçons

2.1.1. Les stéréotypes de genre : définition

Le terme "stéréotype" dérive du grec "stereos" signifiant "solide", et "tupos" signifiant "empreinte". À l'origine, le terme faisait référence à une technique d'imprimerie consistant à créer une empreinte en relief. On comprend mieux alors la définition des stéréotypes dans le domaine des sciences sociales. Selon LEYENS et al. (1994) (p. 24-25), les stéréotypes renvoient à "un ensemble de croyances partagées à propos des caractéristiques personnelles, généralement des traits de personnalité, mais aussi des comportements, propres à un groupe de personnes". Les stéréotypes sociaux sont également des croyances surgénéralisées sur les traits et les compétences des groupes sociaux, au point où les différences interindividuelles vont être systématiquement sous-estimées (JUDD & PARK, 1993).

Plus particulièrement, les stéréotypes de genre font référence aux traits et compétences qui sont attribués différemment aux femmes et aux hommes. Les femmes sont généralement caractérisées par des traits dits de communalité (e.g., gentilles, serviables, chaleureuses, communicatives), tandis que les hommes sont davantage associés à des traits dits d'agentisme (e.g., agressifs, forts, compétitifs, indépendants, ambitieux) (SCZESNY et al., 2018a). Des capacités différentes sont également attribuées aux hommes et aux femmes. Par exemple, les hommes sont considérés comme meilleurs en mathématiques que les femmes, tandis que les femmes sont considérées comme meilleures que les hommes en lettres (ANDRE et al., 1999 ; MAKAROVA et al., 2019). Les stéréotypes de genre sont connus de toutes et tous au sein d'une société, des adultes comme des enfants. En effet, les enfants sont non seulement conscients des identités de genre dès l'âge de 2 ans (HUSTON, 1985), mais ils sont également confrontés très tôt aux stéréotypes de genre (EISENBERG et al., 1996). Les enfants sont exposés aux stéréotypes de genre dès leur plus jeune âge par leur environnement

(parents, voisins, membres de la famille, ...) où les hommes tendent à passer plus de temps au travail alors que les femmes passent plus de temps à effectuer les tâches ménagères (KAN et al., 2011). Aussi, les médias à heure de grande écoute montrent plus fréquemment des hommes dans des situations relatives au travail, et des femmes dans des situations relatives à la famille et au foyer (LAUZEN et al., 2008). Les enfants retrouvent ces mêmes stéréotypes dans leurs jouets, où ceux à destination des filles sont associés à l'attrait physique, à l'éducation et aux compétences domestiques, tandis que les jouets à destination des garçons sont considérés comme violents, compétitifs, excitants et quelque peu dangereux (BLAKEMORE & CENTERS, 2005).

2.1.2. Des perceptions stéréotypées des compétences en mathématiques

Compte tenu de cette exposition quasiment constante aux rôles de genre traditionnels, il n'est finalement pas surprenant que les jeunes enfants eux-mêmes reproduisent les stéréotypes de genre en ayant des croyances stéréotypées en matière de capacités (LEVY et al., 2000), de préférence de jeu (BAKER et al., 2016), de préférence de leurs pairs (SEBANC et al., 2003) et d'aspirations professionnelles dès leur plus jeune âge (WIGFIELD & ECCLES, 2000). SHERMAN et ZURBRIGGEN (2014) ont par exemple montré que, lorsque les jeunes participantes de 4 à 7 ans jouaient avec une poupée Barbie, elles pensaient que les garçons pouvaient faire plus de métiers que les filles. En revanche, les participantes pensaient que les garçons comme les filles pouvaient accéder aux mêmes métiers lorsqu'elles jouaient avec un jouet perçu comme étant plus neutre et qui ne véhiculait pas de stéréotype de genre. De plus, CHAMBERS (1983) et J. STEELE (2003) ont montré dans leurs études que lorsque des enfants d'environ 6 ans devaient dessiner des scientifiques, ils tendaient à dessiner des hommes. Les filles du même âge dans l'étude de BIAN et al. (2017) avaient tendance à moins exprimer que les filles sont "vraiment, très intelligentes" en comparaison aux garçons avec leurs pairs masculins. Ces filles tendaient également à éviter les activités dites pour les enfants "vraiment, très intelligents". Les jeunes filles ont également tendance à penser que les matières où elles sont le moins douées concernent les sciences, alors que les garçons ont tendance à penser qu'ils sont moins doués en lecture (FREEDMAN-DOAN et al., 2000).

D'autres recherches ont illustré des perceptions stéréotypées comparables chez les élèves plus âgés, dont le degré peut varier selon le genre et le comportement de leurs enseignants. L'enquête de SANSONE (2019) a par exemple montré l'existence d'un lien entre les croyances des étudiants sur les capacités des femmes en mathématiques et en sciences, et le genre de leur professeur. Les étudiants, femmes comme hommes, étaient moins enclins à dire que les hommes sont meilleurs que les femmes en mathématiques et en sciences quand leur professeur était une femme. L'auteur a souligné la cohérence de ces résultats avec l'effet positif de femmes leaders sur les aspirations des jeunes femmes (BEAMAN et al., 2012). Non seulement le genre de leur professeur est important dans leur conception des différences de compétences entre

hommes et femmes en mathématiques, mais le comportement de leur professeur l'est également. En effet, une enquête à grande échelle conduite par FORGASZ et LEDER (1996) a montré que les étudiants qui percevaient leur professeur de mathématiques comme étant soucieux/intéressé par eux en tant qu'individus (plutôt qu'en tant que membre d'un groupe de genre) avaient plus tendance à adopter des schémas de pensée rationnelle, utiles pour atteindre leurs objectifs en mathématiques. Cela était d'autant plus vrai pour les filles. Ces résultats se retrouvent dans l'enquête de SANSONE (2019). Autre résultat notable de l'étude de FORGASZ et LEDER (1996) : les professeurs de mathématiques favorisaient les garçons par rapport aux filles. Ce favoritisme se traduisait par plus d'interactions avec les garçons, une plus grande tolérance envers les comportements inappropriés des garçons, et des attentes plus élevées à leur égard. Une méta-analyse conduite par LINDBERG et al. (2010) à partir de 242 études publiées entre 1990 et 2007 a indiqué que même lorsque les filles et les garçons avaient les mêmes résultats en mathématiques, les étudiantes exprimaient plus d'anxiété et de malaise, et moins d'intérêt et de sentiment de réussite en cours de mathématiques que leurs camarades masculins.

Enfin, les perceptions stéréotypées des compétences en mathématiques s'observent aussi chez les parents eux-mêmes qui tendent à expliquer différemment le succès en mathématiques de leurs enfants en fonction de leur genre. Leur fils réussit ainsi en mathématiques grâce à son talent naturel, alors que la réussite de leur fille est due à ses efforts (D. K. YEE & ECCLES, 1988). Ces résultats ont été répliqués dans une autre étude (RÄTY et al., 2002), où les auteurs ont montré que les parents de garçons ont évalué les compétences en mathématiques de leur enfant comme plus grandes que les parents des filles. À l'inverse, ces derniers attribuent plus de compétences en lecture à leurs filles que ne le font les parents de garçons. Les auteurs ont suggéré qu'en justifiant la réussite en mathématiques des filles par leurs efforts, et celle des garçons par leur talent, les parents pourraient susciter des doutes quant à la possibilité de leur fille de pouvoir réussir en mathématiques tout au long de la scolarité. En parallèle, cela pourrait encourager les garçons dans leur avenir en mathématiques (D. K. YEE & ECCLES, 1988). En effet, LESLIE et al. (2015) ont montré que les étudiants pensent qu'ils doivent être naturellement doués dans les mathématiques et les sciences pour y réussir, ce qui peut encourager les garçons compte tenu des explications données à leur réussite en mathématiques par leurs parents. Ces différentes explications selon le genre des élèves ne sont pas bénéfiques pour les femmes et les filles, car le stéréotype de genre associé aux femmes dans le domaine des mathématiques véhicule l'idée qu'elles sont biologiquement inférieures, autrement dit, qu'elles ne sont pas naturellement douées en mathématiques.

2.1.3. Controverse autour de l'idéologie du don en mathématiques

Toutes les différences présentées précédemment reflètent l'influence des stéréotypes de genre et les nourrissent en retours. Ces stéréotypes conduisent les individus

à considérer les hommes comme étant naturellement meilleurs que les femmes en mathématiques et en sciences. L'idée que les hommes possèdent un don inné pour les mathématiques est souvent utilisée pour expliquer la sous-représentation des filles et des femmes dans les filières mathématiques et pourquoi elles obtiennent parfois de moins bonnes performances sur les tests standardisés, tout en excluant d'autres explications possibles. En 2005, le président de l'université de Harvard, Lawrence Henry Summers, a suggéré que la sous-représentation des femmes dans les sciences et l'ingénierie proviendrait de différences biologiques et d'aptitudes intrinsèques qui favorisent les hommes dans ces domaines-là, minimisant le rôle de la socialisation et des discriminations qui ne seraient que des « facteurs moindres » (SUMMERS, 2005). Le débat autour des facteurs explicatifs de la réussite des femmes et des hommes en mathématiques oppose donc principalement deux types de facteurs, biologiques et socio-culturels.

D'un côté, les facteurs biologiques sont mis en avant pour défendre l'idée que l'agentisme et la communalité seraient respectivement associées aux hommes et aux femmes en raison de l'existence de "cerveaux féminins" et de "cerveaux masculins" (COHEN, 2003). Les niveaux de testostérone et d'ocytocine prédisposeraient intrinsèquement les femmes et les hommes à des activités et comportements différents. De même, d'autres auteurs ont suggéré que l'écart en mathématiques en faveur des hommes peut être dû à leur taux plus élevé de testostérone prénatale (BENBOW, 1988; BENBOW & STANLEY, 1980). Cependant, une hormone n'est pas forcément la cause d'un comportement spécifique : la testostérone peut par exemple susciter un comportement prosocial, et pas seulement un comportement agressif ou compétitif (VAN HONK et al., 2011). Aussi, une étude évaluant plus de 2100 cerveaux humains a montré que l'architecture typique d'un cerveau d'un homme et en fait l'architecture typique de celui d'une femme, et vice-versa (JOEL et al., 2018). Enfin, NASH et GROSSI (2007) et SAGUY et al. (2021) ont montré qu'il n'y a pas de résultat probant sur les théories biologisantes sur les différences de comportements et de compétences entre hommes et femmes. En effet, aucune étude expérimentale n'a montré, à ce jour et à notre connaissance, que des différences biologiques peuvent expliquer les écarts de performance en mathématiques en faveur des hommes.

D'un autre côté, d'autres recherches montrent que le comportement des femmes et des hommes est contraint par des "scripts sociaux", appris par la socialisation et la position sociale dans la société, qui conduisent chaque individu à avoir un ensemble d'attentes spécifiques sur le comportement de l'autre personne, en fonction de son genre (EAGLY et al., 2000; RANGEL & KELLER, 2011). Les hommes sont ainsi censés occuper des positions à plus haute responsabilité et être plus compétents que les femmes dans les domaines scientifiques. De plus, un grand nombre d'études (nous en avons déjà évoqué certains, et allons présenter d'autres études plus en détail à la suite de ce chapitre) défendant l'idée d'une origine sociale et culturelle des différences de performances des femmes et des hommes montrent que les stéréotypes de genre jouent un rôle majeur sur les écarts en mathématiques en faveur des hommes. Il est

particulièrement intéressant de noter que ces études montrent que le rôle de ces stéréotypes dans les écarts de performance en mathématiques dépend fortement du contexte situationnel et non dispositionnel, et est sensible aux influences extérieures. La recherche montre qu'il est alors possible de réduire voire d'annuler les écarts de performance en mathématiques en faveur des hommes. Ces recherches tendent à montrer que la biologie n'est pas suffisante pour expliquer cet écart de performance en mathématiques. Dans cette thèse, nous nous intéressons plus particulièrement aux études expérimentales sur les effets des stéréotypes de genre sur la performance en mathématiques des filles et des femmes (Section 2.3), à travers le phénomène dit de menace du stéréotype que nous présentons dans la prochaine section.

2.2. Le phénomène de menace du stéréotype : définition et principaux mécanismes explicatifs

Le phénomène appelé "menace du stéréotype" correspond à la crainte d'être négativement stéréotypé lors d'un test de compétence ou de confirmer un stéréotype négatif envers le groupe social auquel on appartient, ce qui peut conduire à des contre-performances. Ce phénomène a été mis en évidence pour la première fois par C. M. STEELE et ARONSON (1995) concernant le stéréotype selon lequel les Afro-Américains auraient un niveau d'intelligence inférieur à celui des Blancs. Les auteurs ont testé un nouveau facteur explicatif des écarts de performance aux tests d'intelligence verbale (e.g., des tests de quotient intellectuel) entre les Afro-Américains et les Blancs en faveur de ces derniers. Ils ont émis l'hypothèse que le simple fait de connaître le stéréotype négatif, sans forcément y adhérer, réduirait les performances des Afro-Américains sur les tests d'intelligence verbale. Les enjeux associés à cette hypothèse sont considérables, d'une part en raison de l'utilisation de ces tests aux États-Unis pour l'admission à l'Université, et d'autre part en raison de l'explication prédominante de la moindre réussite académique des Afro-Américains par leur supposée infériorité intellectuelle. C. M. STEELE et ARONSON (1995) ont montré que les participants Afro-Américains obtenaient de moins bons résultats que les participants Blancs à un test difficile standardisé d'intelligence verbale lorsque le test était décrit comme un test d'intelligence (consigne standard de présentation du test lors de l'examen d'entrée à l'Université). En revanche, les résultats des participants Noirs et Blancs étaient tout aussi élevés lorsque le même test était présenté comme un test de résolution de problèmes, c'est-à-dire sans mentionner l'évaluation des capacités intellectuelles. Par ailleurs, les Afro-Américains dans cette condition réussissaient mieux le test que leurs homologues ayant passé le test dans des conditions ordinaires. La publication de C. M. STEELE et ARONSON (1995) comprend quatre expérimentations qui conduisent toutes aux mêmes résultats. Présenter le test de manière à atténuer la pertinence du stéréotype négatif envers les Afro-Américains permet de les soulager de la crainte de confirmer les attentes négatives véhiculées par ce stéréotype envers leurs capacités intellectuelles. Comme nous l'avons dit plus haut, il est important de noter que décrire le test comme un test d'intelligence verbal correspond à la description habituelle

(instruction standard) de ce type de test pour l'admission à l'Université aux États-Unis. Par conséquent, la condition de menace du stéréotype correspond à la condition classique lorsque les étudiants passent un examen : la menace du stéréotype se trouve, par défaut, dans la situation de test elle-même, comme l'énonce le titre de l'article de C. M. STEELE (1997) ("*la menace est dans l'air*").

Dans un contexte où les écarts de performance entre Blancs et Afro-Américains aux tests d'intelligence verbale étaient justifiés par des différences intrinsèques d'intelligence en faveur des Blancs, C. M. STEELE et ARONSON (1995) ont ainsi déstabilisé ces explications biologisantes pour la première fois. Ces auteurs ont tout de même précisé qu'on ne peut pas restreindre les différences de performances entre groupes sociaux aux seuls effets des stéréotypes, les causes étant évidemment plurielles (C. M. STEELE & ARONSON, 2004). Depuis 1995, de nombreuses études sur les effets de menace du stéréotype ont été publiées dans le monde, comme en atteste la simple requête "*stereotype threat*" (e.g., "menace du stéréotype") sur la base de données Web of Science : en cherchant dans les titres, les résumés, les mots clés et les mots clés additionnels, la requête a sélectionné 3488 articles. Les chercheurs ont étendu les résultats de C. M. STEELE et ARONSON (1995) à d'autres groupes sociaux avec leurs stéréotypes associés. Ainsi, ils ont observé les effets de menace du stéréotype chez les élèves d'origine sociale défavorisée (CROIZET & CLAIRE, 1998) ou encore chez les jeunes français d'origine maghrébine (CHATEIGNIER et al., 2009) aux tests intellectuels. Des recherches ont également montré que la menace du stéréotype agit sur les garçons sur les tests de lecture (PANSU et al., 2016), les personnes âgées sur les tests de mémoire (GAUTHIER et al., 2020), et les filles et les femmes sur les tests de mathématiques (RÉGNER et al., 2014 ; SPENCER et al., 1999). Les groupes sociaux typiquement avantagés dans nos sociétés peuvent également faire face à la menace du stéréotype, comme montré par ARONSON et al. (1999). Les participants, des hommes étudiants Blancs et Américains, ont moins bien réussi en mathématiques lorsqu'il a été dit que leurs résultats seraient comparés avec ceux d'étudiants Asiatiques. Ils ont ainsi été confronté au stéréotype opposant leurs compétences en mathématiques avec celles des personnes Asiatiques, supposées être bien meilleures. Toutefois, les études montrent surtout que les stéréotypes sociaux induisent un effet de menace du stéréotype chez les groupes stigmatisés et au contraire un effet boost (ou *stereotype lift*) chez les groupes habituellement avantagés (WALTON & COHEN, 2003). En effet, dans la mesure où les stéréotypes tendent à opposer deux groupes sociaux en dévalorisant l'un au profit de l'autre, l'activation de ces stéréotypes lors d'une situation de test place le groupe avantagé dans une position comparative psychologiquement avantageuse et confortable, susceptible d'améliorer leurs performances (DANSO & ESSES, 2001).

Les personnes confrontées à la menace du stéréotype sont motivées à infirmer les stéréotypes négatifs ciblant leur identité sociale (NUSSBAUM & STEELE, 2007 ; VANDELLO et al., 2008), ou au moins à éviter de les confirmer (BRODISH & DEVINE, 2009 ; CHALABAEV et al., 2012). Ces motivations représentent une pression supplémentaire à la réussite (ou à l'évitement de l'échec) à laquelle les individus non stigmatisés

ne sont pas confrontés, et peuvent entraîner une contre-performance. En se fondant sur la théorie de la facilitation sociale de ZAJONC (1965), ces motivations pousseraient à donner une réponse dominante lors d'une tâche (HARKINS, 2006). Sur une tâche facile, la réponse dominante est généralement correcte, alors qu'elle est souvent incorrecte lors de tâches difficiles. Cette pression à réussir peut également nuire aux performances par la perturbation des systèmes cognitifs nécessaires à la résolution de problèmes, en demandant notamment plus d'effort à l'individu concernant la régulation de ses émotions et de ses pensées parasites. SCHMADER et al. (2008) ont soutenu que la menace du stéréotype active une réponse physiologique au stress, un surcontrôle cognitif, et des tentatives de suppression des pensées intrusives négatives. Ces trois mécanismes consomment les capacités de la mémoire de travail, une ressource cognitive limitée, qui sert notamment à réguler, contrôler, et stocker des informations temporaires présentant un intérêt immédiat pour résoudre une tâche cognitive (MIYAKE, 1999). BEILOCK et al. (2006), BRODISH et DEVINE (2009) et SCHMADER et al. (2008) ont en effet montré un surcontrôle cognitif de la part des personnes ciblées par la menace du stéréotype, qui concentrent leur attention sur elles-mêmes et sur leurs performances, et deviennent très (trop) vigilantes pour détecter leurs éventuelles erreurs (e.g., relire constamment ses réponses pour un exercice en particulier). De plus, des études ont montré que lors de situations de menace du stéréotype, les individus ciblés peuvent douter d'eux-mêmes (C. M. STEELE & ARONSON, 1995), se décourager (KELLER & DAUENHEIMER, 2003), être inquiets par rapport à la tâche (BEILOCK et al., 2007), et avoir des pensées négatives liées au stéréotype (LOGEL, ISERMAN et al., 2009). Ces pensées et émotions négatives doivent être inhibées pour se concentrer sur la tâche en cours. Or, supprimer une pensée est un processus complexe : WEGNER (1994) a suggéré qu'il faudrait consciemment éviter d'y penser, tout en surveillant son environnement pour vérifier la présence de ladite pensée. Enfin, d'autres études ont montré que la menace du stéréotype engendre une réponse physiologique au stress qui active les zones neuronales associées aux traitements des informations sociales et à la régulation des émotions, plutôt que les zones utiles à la résolution de problèmes (KRENDL et al., 2008 ; WRAGA et al., 2007). Le stress de la menace du stéréotype peut également activer le système nerveux sympathique (M. C. MURPHY et al., 2007), et augmenter la pression sanguine (BLASCOVICH et al., 2001) et le rythme cardiaque (MENDES et al., 2002).

La menace du stéréotype engendre donc un stress supplémentaire lié au stéréotype, en plus du stress lié à la peur de l'échec que tout individu peut éprouver face à un test difficile, dans une situation évaluative à fort enjeu. Les efforts de la mémoire de travail pour mettre de côté toutes ces émotions négatives entraînent une diminution de sa capacité (BEILOCK et al., 2007). La difficulté à se concentrer mentalement qui en résulte (C. M. STEELE, 2011) peut conduire à persister dans une stratégie de résolution non-pertinente ou non-efficace (CARR & STEELE, 2009). Si la situation devient chronique, cela peut entraîner une perte de motivation et d'intérêt pour le domaine (C. M. STEELE, 1997), une mauvaise estime de soi, et un isolement dans les études plutôt que demander de l'aide aux autres (C. M. STEELE, 2011). Elle peut également

expliquer, du moins en partie, pourquoi les femmes sont moins nombreuses que les hommes à poursuivre une carrière scientifique (BEN-ZEEV, DUNCAN & FORBES, 2005), ainsi que l'écart de performance entre hommes et femmes en mathématiques dans les tests standardisés difficiles.

2.3. Les effets de menace du stéréotype sur la performance des filles et des femmes en mathématiques

2.3.1. Illustration de la menace du stéréotype en mathématiques

À la suite des travaux précurseurs de C. M. STEELE et ARONSON (1995), de nombreuses recherches ont étudié les effets de menace du stéréotype sur la performance des filles et des femmes en mathématiques. SPENCER et al. (1999) ont été les premiers à tester ces effets en mathématiques, auprès d'une population d'étudiantes et d'étudiants, en utilisant un test standardisé. Après avoir montré dans une première étude que les écarts de performance entre hommes et femmes s'observent uniquement sur les exercices complexes du test de mathématiques, ces auteurs ont réalisé deux études pour tester l'effet de menace du stéréotype sur ces exercices difficiles. Pour rappel, C. M. STEELE et ARONSON (1995) avaient manipulé la "diagnosticité" du test pour induire (i.e., test d'intelligence) *versus* réduire (i.e., test de résolution de problème) l'effet de menace du stéréotype. En revanche, SPENCER et al. (1999) ont choisi de maintenir la présentation du test comme un test de mathématiques dans toutes les conditions, et ont plutôt évoqué la présence *versus* l'absence de différences de performance entre hommes et femmes sur le test pour, respectivement, induire *versus* réduire l'effet de menace du stéréotype. Plus précisément, dans leur étude n°2, le même test de mathématiques était décrit soit comme ayant révélé par le passé des disparités de performance entre les hommes et les femmes (condition menaçante), soit comme n'ayant jamais révélé d'écart de performance entre hommes et femmes. Cette dernière consigne, appelée par les auteurs "consigne de falsification", avait pour objectif de rendre le stéréotype de genre non pertinent pour la situation de test, et soulager ainsi les femmes du stress afférent. Les résultats ont montré l'écart habituel de performance en faveur des hommes sur ce test standardisé de mathématiques lorsque les différences de genre étaient évoquées. Même si le sens de ces différences n'était pas précisé, SPENCER et al. (1999) considèrent que cette consigne a rendu le stéréotype de genre saillant dans la situation de test. En revanche, les femmes ont obtenu des performances tout aussi élevées que celles des hommes lorsque le test était accompagné de la consigne de falsification. Dans une troisième étude, SPENCER et al. (1999) ont montré qu'il n'était pas nécessaire de mentionner explicitement le stéréotype pour observer les effets de menace du stéréotype. Ils ont comparé la condition de falsification similaire à celle de l'étude précédente à une condition contrôle où aucune autre

information n'était fournie sur le test de mathématiques. Autrement dit, la condition contrôle correspondait aux conditions normales et standard d'administration d'un test de mathématiques dans un contexte d'examen. Alors que les femmes ont moins bien réussi le test de mathématiques que les hommes dans la condition contrôle, elles ont obtenu des performances tout aussi élevées qu'eux lorsque la consigne de falsification était ajoutée.

SPENCER et al. (1999) ont ainsi répliqué les résultats de C. M. STEELE et ARONSON (1995) chez les femmes en mathématiques, en montrant que la menace du stéréotype joue un rôle dans les écarts de performance en mathématiques en faveur des hommes. Ainsi, pour la première fois, des études expérimentales ont rendu les explications biologisantes des différences genrées en mathématiques difficiles à soutenir. En rendant le stéréotype de genre en mathématiques temporairement non pertinent, la consigne de falsification a permis aux participantes de mieux réussir le test, et tout aussi bien que leurs homologues masculins. Rappelons que la consigne qui mentionne l'existence de différences de performance entre hommes et femmes sur ce test (étude 2), ne précise nullement en faveur de quel groupe ces différences étaient orientées. Selon SPENCER et al. (1999), on peut supposer que le stéréotype est si fort que les participants, hommes comme femmes, ont probablement inféré que les hommes avaient obtenu les meilleurs résultats. L'étude 3 a permis de montrer que le simple fait de présenter un test comme évaluant les compétences en mathématiques suffit pour déclencher un effet de menace du stéréotype chez les étudiantes. En partant du principe que le test évalue leurs compétences en mathématiques, les femmes se retrouvent ainsi dans une situation dans laquelle elles pourraient être jugées sous le prisme du stéréotype négatif qui suppose que les femmes sont naturellement inférieures aux hommes en mathématiques. D'autres études ont confirmé que les stéréotypes activés explicitement (e.g., test de mathématiques + « des différences de performance entre hommes et femmes ont été observées sur ce test ») et implicitement (e.g., test de mathématiques, sans information supplémentaire) sont tout aussi nocifs pour la performance des femmes en mathématiques (CAMPBELL & COLLAER, 2009 ; SMITH & WHITE, 2002).

Depuis, de nombreuses études ont répliqué l'effet de menace du stéréotype chez les femmes en mathématiques (AMBADY et al., 2004 ; KAO et al., 2024 ; SCHMADER & JOHNS, 2003 ; VAN LOO et al., 2013), et également chez les filles (AMBADY et al., 2001 ; GALDI et al., 2014 ; PICHO & STEPHENS, 2012 ; RÉGNER et al., 2014 ; TOMASETTO et al., 2011). Les études de HUGUET et REGNER (2007) et HUGUET et RÉGNER (2009) ont montré que les effets négatifs de menace du stéréotype étaient présents même lorsque des élèves de 6e et de 5e pensaient passer un test de mathématiques, alors qu'il s'agissait d'un test de mémoire où ils devaient dessiner une figure. Lorsque le test a été présenté comme un test de dessin, les filles ont obtenu de meilleures performances que les garçons, alors que le contraire s'est produit lorsque le test a été décrit comme un test de géométrie. Les auteurs ont donc répliqué les effets de menace du stéréotype d'une

part en condition quasi-ordinaire [1] de classe, et d'autre part lorsque les filles pensaient que le test évaluait les capacités en mathématiques.

Ces résultats permettent de comprendre qu'une des particularités de la menace du stéréotype est qu'aucune intervention n'est nécessaire de la part des chercheurs pour la faire apparaître : une situation d'évaluation sur un test difficile de mathématiques représente, par défaut, une situation de menace de stéréotype pour les filles et les femmes dans ce domaine de compétence. À l'inverse, une intervention est nécessaire pour réduire voire annuler les effets de menace du stéréotype (e.g., présenter le test comme mesurant autre chose que la compétence ciblée par le stéréotype ; ajouter la consigne de falsification). Cette distinction est fondamentale pour éviter un biais méthodologique fréquent qui consiste à supposer qu'une condition de menace du stéréotype requiert la mise en place d'une pression supplémentaire qui n'existe pas lors des tests dans la "vraie" vie, et que les instructions de test standard peuvent être utilisées pour opérationnaliser une condition de contrôle sans menace du stéréotype (MAZEROLLE et al., 2020 ; C. M. STEELE & DAVIES, 2003). Au contraire, dans les études expérimentales, le groupe contrôle qui consiste à simplement présenter le test pour ce qu'il est correspond en fait à une condition de menace du stéréotype. Pour tester les effets de menace du stéréotype, il convient donc d'être vigilant pour s'assurer de comparer au moins deux conditions (C. M. STEELE & ARONSON, 2004) :

- **Une condition de menace du stéréotype** : il peut s'agir d'une situation classique d'évaluation en mathématiques (e.g., un contrôle de mathématiques noté), où le stéréotype est activé de façon implicite. Il peut s'agir également d'une situation de test moins standard où le stéréotype est rendu plus explicite (e.g., mentionner des écarts de performances entre hommes et femmes en mathématiques ; demander aux élèves d'indiquer le sexe avant de faire le test).
- **Une condition de menace réduite** : c'est ici que les chercheurs sont actifs, et testent des méthodes pour réduire voire supprimer les effets de menace du stéréotype. Nous avons présenté une de ces méthodes ci-dessus, en décrivant un test de mathématiques comme n'ayant jamais montré de différence de performance entre hommes et femmes (i.e., consigne de falsification). Nous présentons d'autres méthodes en Section 2.3.4.

D'autres caractéristiques doivent être prises en compte lorsque les chercheurs testent la menace du stéréotype, car elles peuvent altérer les résultats. Ces facteurs, appelés modérateurs, peuvent par exemple être liées aux profils des participantes. Nous présentons ces caractéristiques dans la prochaine sous-section.

1. Les études sur la menace du stéréotype dans les classes sont en situation quasi-ordinaire car la présence d'un expérimentateur n'est pas habituelle.

2.3.2. Modérateurs des effets de menace du stéréotype

En statistiques, les variables modératrices peuvent influencer la force d'une relation entre les variables indépendantes et dépendantes, en la renforçant ou en l'affaiblissant. Les différences individuelles peuvent modérer les effets de menace du stéréotype sur la performance de deux manières : d'une part, en influant sur le degré d'investissement des femmes dans la situation de test, et d'autre part, en influençant leur capacité à faire face à une pression supplémentaire (SPENCER et al., 2016). Identifier les modérateurs de la menace du stéréotype permet de mieux comprendre à la fois les processus impliqués dans ce phénomène et les résultats obtenus dans les études. Nous présentons ci-dessous une synthèse des différents modérateurs susceptibles d'altérer les effets de menace du stéréotype.

2.3.2.1. Identification aux mathématiques

Des recherches ont montré que le degré d'identification des femmes aux mathématiques impacte les effets de menace du stéréotype. En effet, la menace du stéréotype affecte plus fortement les femmes qui ont les compétences et la confiance en elles requises pour s'identifier aux mathématiques (C. M. STEELE, 1997). Les mathématiques peuvent être perçues comme importantes et occuper une position centrale dans la vie d'un individu. Les personnes ayant du mal à s'identifier avec le domaine risquent d'obtenir de faibles résultats, que la menace du stéréotype soit présente ou non, car la réussite en mathématiques n'est pas importante pour elles. En revanche, pour les femmes s'identifiant fortement aux mathématiques, l'enjeu que représente un test de mathématiques est bien plus important. Échouer à un test de mathématiques pourrait réduire leur confiance en elles ainsi que leur image, et pourrait les conduire à exclure les mathématiques de leur conception de soi pour protéger leur estime de soi (SCHMADER et al., 2004). LESKO et CORPUS (2006) ont noté que ce processus de désidentification est cohérent avec la littérature sur l'identité sociale, suggérant que des changements dans les valeurs personnelles se produisent souvent lorsque les individus perçoivent que la réussite est improbable dans un domaine central de leur identité (CROCKER & MAJOR, 1989 ; MAJOR et al., 1998). Dans une étude longitudinale, M. J. FISCHER (2010) a montré que se désidentifier de n'importe quel domaine académique entraîne des contre-performances dans le dit-domaine. C'est ainsi que les effets de menace du stéréotype ont été principalement examinés et constatés chez les femmes très performantes ou s'identifiant au domaine des mathématiques (BEN-ZEEV, FEIN & INZLICHT, 2005 ; BROWN & PINEL, 2003 ; LESKO & CORPUS, 2006 ; MARTENS et al., 2006). KELLER (2002) a par exemple supposé qu'il n'a pas répliqué les résultats habituels de la menace du stéréotype parce qu'il n'a fait aucune présélection des participantes selon leur niveau d'identification en mathématiques. Dans son étude en situation de classe quasi-ordinaire, les femmes et les hommes ont obtenu les mêmes résultats dans la condition où il a été dit que le test de mathématiques évaluait les compétences individuelles en mathématiques, alors qu'il s'agissait d'une situation de menace du stéréotype. En revanche, les femmes ont moins bien réussi

que les hommes lorsque le stéréotype a été explicitement énoncé en précisant que les hommes obtenaient de bien meilleurs résultats que les femmes dans ce test de mathématiques. L'auteur a alors suggéré que lorsque les participantes ne sont pas identifiées aux mathématiques, une référence explicite au stéréotype négatif pourrait être nécessaire pour déclencher les mécanismes de la menace du stéréotype. Ceci pourrait expliquer, au moins en partie, pourquoi certaines études ne répliquent pas les résultats sur la menace du stéréotype (STRICKER & BEJAR, 2004). Pour mesurer l'identification au domaine, les chercheurs peuvent demander aux participants d'indiquer l'importance des mathématiques pour l'image qu'ils ont d'eux-mêmes à l'aide de questionnaire (e.g., "Il est important pour moi d'être bon en mathématiques.") sur une échelle de Likert (KELLER, 2007). Dans cette thèse, nous évaluons l'importance accordée aux mathématiques ainsi que le sentiment de réussite dans ce domaine.

2.3.2.2. Identification à son groupe de genre

Comme dit précédemment, pour se protéger le plus possible de la menace du stéréotype, les femmes peuvent se désidentifier du domaine. Une autre stratégie, notamment des femmes dans les filières scientifiques, est de se distancier de leur groupe de genre ou de leur identité de genre pour s'adapter à la masculinité de leur environnement de travail (DERKS et al., 2016; FANIKO et al., 2021; von HIPPEL et al., 2011). Ainsi, dans la deuxième étude de WOUT et al. (2008), lorsque la menace ciblait les femmes en tant que groupe social (e.g., en précisant que le score moyen des femmes participantes serait comparé à celui des hommes participants), plus les participantes s'identifiaient aux femmes, moins elles répondaient correctement aux questions. Plus particulièrement, les étudiantes en psychologie dans l'étude de CRISP et al. (2009) ont obtenu de moins bons résultats en situation de menace du stéréotype. En revanche, un boost du stéréotype a été observé pour celles en ingénierie. Les auteurs ont expliqué cela par le fait que les ingénieures de cette étude s'identifiaient plus à l'ingénierie qu'à leur identité de genre, alors que les étudiantes en psychologie s'identifiaient plus à leur genre qu'à leur filière. La force de l'identification au genre *féminin* modère ainsi les effets de menace du stéréotype. L'identification au groupe *femme* ou *fille* peut être mesurée implicitement via des tests comme le TAI (Test des Associations Implicites, GREENWALD et al., 2003) permettant d'évaluer si les individus ont tendance à s'associer plus fréquemment aux hommes/garçons, aux femmes/filles, ou à aucun de ces groupes (CVENCEK et al., 2015). Elle peut également être évaluée explicitement via un questionnaire (e.g., "Être une femme/un homme est une part importante de mon égo.") (SCHMADER, 2002).

2.3.2.3. Associations implicites stéréotypées

La force des associations implicites en mémoire sémantique entre masculin-sciences et féminin-lettres peut également modérer la menace du stéréotype. Les associations implicites peuvent être mesurées à l'aide de test comme le TAI permettant de déterminer si les individus ont tendance à associer plus fréquemment les hommes aux

sciences et les femmes aux lettres plutôt que l'inverse (GREENWALD et al., 2003). Les associations implicites d'un individu ne révèlent pas nécessairement ses opinions véritables. Le score du TAI peut même contredire les croyances explicites d'une personne. Le test vise plutôt à estimer la force des associations entre les informations stockées dans la mémoire sémantique. Chez les adultes (NOSEK et al., 2009) comme chez les enfants (CVENCEK et al., 2011), les individus ont en général des associations plus fortes entre masculin-sciences qu'entre féminin-sciences, attestant de la présence et de la force des stéréotypes de genre dans nos quotidiens.

NOSEK et al. (2002) ont montré que pour les participantes, plus les automatismes stéréotypés étaient forts, moins elles aimaient les mathématiques, moins elles s'identifiaient à ce domaine, et leur performance en mathématiques étaient moins bonnes. L'inverse se produisait pour les hommes participants : les automatismes stéréotypés étaient positivement liés à leur identification au domaine des mathématiques et à leur performance. Les résultats de KIEFER et SEKAQUAPTEWA (2007) suggèrent que seules les femmes ayant un faible score au TAI bénéficient de la phrase de falsification (ici en précisant que le test n'évalue pas les capacités en mathématiques). Pour autant, FRANCESCHINI et al. (2014) ont trouvé des résultats inverses : dans une condition de menace réduite (ici en précisant que les femmes ont obtenu de meilleurs résultats que les hommes), c'étaient les femmes avec des automatismes stéréotypés élevés qui produisaient de meilleures performances. Ces résultats contradictoires peuvent être dûs aux consignes utilisées dans le but de réduire la menace. Il est probable que pour les femmes ayant un score faible au TAI, préciser que l'on n'évalue pas les compétences en mathématiques est suffisant pour augmenter leurs résultats. En revanche, une consigne plus forte mettant en doute le stéréotype (e.g., "les femmes ont mieux réussi que les hommes") serait nécessaire pour celles ayant un score au TAI plus élevé. Enfin, GALDI et al. (2014) ont trouvé que ces automatismes avaient un rôle de médiateur plutôt que de modérateur, uniquement chez les filles d'environ 6 ans. GALDI et al. (2014) ont de plus montré qu'à cet âge-là, la stéréotypie implicite était malléable chez les filles. Elles avaient en effet obtenu un score plus faible au TAI après avoir colorié une image où une fille réussit un exercice de mathématiques et un garçon échoue, comparativement aux autres filles ayant colorié une image où c'est le garçon qui réussit et la fille qui échoue en mathématiques, et aux filles ayant colorié un paysage.

Ces études contiennent des résultats contradictoires, cependant explicables par la différence de procédure pour induire une situation de menace réduite. Dans cette thèse, nous avions donc évalué l'effet modérateur du TAI sur la performance en mathématiques, l'importance accordée aux mathématiques, ainsi que le sentiment de réussite dans cette discipline. La question reste en effet entière sur ce qui peut se passer chez des élèves de 6e et de 5e. Pour cela, nous avons recodé le TAI genre-mathématiques-lecture pour enfants (CVENCEK et al., 2011), cette fois-ci en français.

2.3.2.4. Hostilité perçue de l'environnement

Les femmes ont connaissance du stéréotype négatif envers leurs capacités en mathématiques, ce qui créé un environnement hostile pour elles ainsi qu'un sentiment de non-appartenance au domaine (GOOD et al., 2012 ; SHAFFER et al., 2013). De plus, l'étude de STOKES et al. (1995) a montré que lorsque les femmes trouvent un environnement de travail non hostile, elles ont plus de chances de conserver leur emploi. L'hostilité de l'environnement joue un rôle dans la sous-représentation des femmes en sciences, mais également dans les effets de menace du stéréotype. LOGEL, WALTON et al. (2009) ont montré dans deux études que des étudiantes en ingénierie ayant interagi avec des hommes présentant des comportements sexistes (e.g., être très proches d'elles, regarder leur corps) ont obtenu de moins bons résultats à un test d'ingénierie comparativement aux étudiantes ayant interagi avec des hommes sans comportements sexistes. Dans une autre étude, les auteurs ont répliqué ces résultats, et ont également montré que ces contre-performances ne se produisaient pas sur un test d'anglais, un domaine dans lequel les femmes ne sont pas stéréotypées négativement. L'hostilité (voire le sexisme) des hommes à leur égard renforce les effets négatifs de menace du stéréotype. D'autres études montrent que la présence d'hommes, même lorsqu'ils ne montrent pas de comportements particuliers, renforce également la menace du stéréotype. PICHO et STEPHENS (2012) ont montré, en situation de classe quasi-ordinaire dans les lycées mixtes d'Ouganda, que les filles en situation de menace du stéréotype (e.g., en précisant que le test a toujours montré des différences de performance entre les filles et les garçons) ont obtenu de moins bons résultats que celles en situation de menace réduite (e.g., un test de mathématiques non noté). Cependant, dans les lycées non-mixtes, les filles ont obtenu les mêmes résultats dans les deux conditions ; elles ont également obtenu de meilleur résultat que celles étant dans un lycée mixte. Ces résultats montrent tout d'abord que la menace du stéréotype est également présente dans les pays non-occidentaux, mais également que la présence de garçons ou d'hommes peut influencer la performance des filles et des femmes. D'autres études vont dans ce sens : dans la deuxième étude de HUGUET et REGNER (2007) en classe chez les collégiens, la condition de menace du stéréotype n'a pas eu d'impact sur les filles lorsque les passations des tests étaient effectuées en groupes non-mixtes, contrairement aux filles dans des groupes mixtes. Aussi, dans une autre étude où les participantes étaient dans un groupe de joueurs représentés par des avatars, la menace du stéréotype agissait davantage lorsqu'elles étaient placées dans un groupe où les femmes étaient minoritaires (INZLICHT & BEN-ZEEV, 2000 ; J.-E. R. LEE & NASS, 2012). En dehors de l'environnement où se déroulent les études, l'entourage des filles peut également avoir un impact sur la menace du stéréotype : TOMASETTO et al. (2011) ont ainsi montré que, parmi leurs participantes en maternelle, la menace du stéréotype impactait plus les jeunes filles dont la mère adhérait au stéréotype de genre en mathématiques.

Il est important de souligner que ces résultats ne doivent pas être interprétés comme une démonstration des avantages des écoles non-mixtes par rapport aux écoles mixtes

(HUGUET & REGNER, 2007). Regrouper des individus selon un groupe social génère voire renforce les stéréotypes et préjugés envers les autres groupes (M. R. CADINU & ROTHBART, 1996; GAERTNER & INSKO, 2000). De plus, la méta-analyse de PAHLKE et al. (2014) sur 184 études a montré que les écoles non-mixtes ne sont pas plus efficaces que les écoles mixtes pour éduquer les filles et les garçons, que ce soit sur leurs performances en sciences ou les stéréotypes de genre. Enfin, l'étude de OSWALD et HARVEY (2000) a montré qu'il est tout de même possible de réduire les effets de menace du stéréotype dans un environnement hostile. Les participantes en condition de menace ont moins bien réussi lorsque l'environnement était hostile pour elle par la présence d'une affiche illustrant une fille qui avait du mal à résoudre un problème de mathématiques facile et un garçon qui résolvait un problème de mathématiques plus complexe. La présence de cette affiche n'a eu aucune incidence sur les participantes en condition de menace réduite, où il était précisé que les hommes et les femmes réussissaient aussi bien dans le test de mathématiques à accomplir. Enfin, l'hostilité peut être réduite, notamment en créant des environnements plus sécurisants pour les filles et les femmes (voir Section 2.3.4.3).

2.3.2.5. Capacité en mémoire de travail

Nous avons présenté en Section 2.2 l'impact de la menace du stéréotype sur la mémoire de travail. En essayant de mettre de côté les émotions négatives générées par la menace du stéréotype, la capacité de la mémoire de travail se réduit, laissant moins de ressources à l'individu pour traiter la tâche à résoudre. SCHMADER et al. (2008) ont émis l'hypothèse que les femmes avec une plus grande capacité dispositionnelle de mémoire de travail disposeraient de plus de ressources pour faire face à la menace du stéréotype, et pourraient alors être moins impactées par les effets négatifs du stéréotype. La capacité de la mémoire de travail, qui peut varier d'un individu à l'autre, peut être mesurée à l'aide de tâche comme le *reading span test* (DANEMAN & CARPENTER, 1980). Ce test consiste à lire à haute voix des séries de phrases, décider si chaque phrase a un sens, mémoriser une lettre, et enfin se rappeler des lettres vues. Le score de la capacité de la mémoire de travail est égal à la proportion moyenne de lettres correctement rappelées dans une série. Dans une étude réalisée auprès d'étudiantes et d'étudiants inscrits en Grandes Ecoles d'Ingénieurs, RÉGNER et al. (2010) ont montré que les différences interindividuelles de capacité de mémoire de travail modèrent les effets de menace du stéréotype. Alors que l'effet classique de menace du stéréotype est observé chez les participants ayant une faible capacité en mémoire de travail, les femmes ayant une plus forte capacité en mémoire de travail ont résisté à cet effet. Cette étude a été réalisée avec des participants ayant de très bons résultats scolaires dans des écoles d'ingénieur hautement sélectives, pour s'assurer que les individus étaient fortement identifiés aux mathématiques et l'ingénierie. Ces résultats montrent que, bien que la menace du stéréotype agisse plus particulièrement sur les femmes fortement identifiées au domaine, ces femmes peuvent tout de même résister aux effets de menace du stéréotype si elles disposent d'une capacité de mémoire de travail suffisamment grande.

2.3.2.6. Vulnérabilité à la pression

Les différences individuelles qui affectent la vulnérabilité des individus face au stress et à la pression peuvent affecter leurs performances en situation de menace du stéréotype. Dans les deux études de M. CADINU et al. (2006), les femmes attribuant les causes d'un événement à leurs propres actions (e.g., « c'est de ma faute si j'ai eu une mauvaise note ») ont obtenu des résultats inférieurs par rapport aux femmes associant les causes d'un événement à des facteurs externes (e.g., « j'ai eu une mauvaise note parce que le professeur est sévère »), mais uniquement en situation de menace du stéréotype. De fortes capacités d'adaptation sont associées à une plus forte résistance face à la menace du stéréotype, comme montré par FORD et al. (2004). Dans leur première étude, les femmes n'utilisant pas l'humour pour faire face aux situations stressantes sous-évaluaient leur performance en mathématiques, comparativement aux hommes et aux femmes utilisant l'humour dans ce genre de situation. Les auteurs ont alors montré dans leur deuxième étude que, en situation de menace du stéréotype, faire preuve d'humour pour mieux gérer les situations stressantes était positivement corrélé aux performances des femmes. De plus, les femmes utilisant l'humour comme stratégie de défense ressentaient moins d'anxiété lors du test. Aussi, dans l'étude de PERRY et SKITKA (2009), les femmes ayant des attentes modestes quant à leurs résultats ont finalement obtenu de meilleurs résultats en situation de menace du stéréotype. À l'inverse, les femmes ayant des attentes élevées ont mieux réussi en situation de menace réduite, où il a été précisé que les femmes et les hommes obtenaient les mêmes résultats sur ce test. Il a déjà été montré qu'en situation de stress, les individus utilisant des pensées pessimistes pour mieux gérer cette pression ne produisent pas de contre-performance s'ils ont la possibilité d'exprimer leurs pensées, et obtiennent d'aussi bons résultats que les individus utilisant les pensées positives pour faire face au stress (NOREM & CANTOR, 1986).

2.3.2.7. Difficulté des exercices

Plusieurs études ont montré que l'écart de performance entre femmes et hommes en faveur de ces derniers s'observe sur les exercices de mathématiques difficiles standardisés (SPENCER et al., 1999). On ne retrouve pas cet écart sur les exercices faciles standardisés. En effet, si la mémoire de travail possède suffisamment de ressources, les femmes peuvent résister aux effets de menace du stéréotype. Puisqu'une tâche facile consomme peu de ressources en mémoire de travail, les femmes peuvent résoudre ce type de tâche malgré leur ressources cognitives diminuées par la menace du stéréotype (O'BRIEN & CRANDALL, 2003). Cependant, si elles ne disposent pas de ressources suffisantes pour à la fois supprimer les pensées négatives et résoudre le test de mathématiques (e.g., lorsque le test est difficile), la menace du stéréotype continuera d'agir, ce qui entraînera une contre-performance (SPENCER et al., 1999). Lorsque les exercices sont trop compliqués, on peut observer un effet plafond où personne ne réussit, hommes comme femmes (SMITH & WHITE, 2002). D'autres études ont montré que la menace du stéréotype a plus d'impact sur les problèmes de mathématiques

verbaux (e.g., avec un énoncé) que les problèmes de mathématiques numériques (e.g., les équations) à cause de l'étape supplémentaire qui demande à élaborer des stratégies pour convertir les énoncés en problèmes numériques. Les participantes de l'étude de QUINN et SPENCER (2001) ont moins bien réussi que les hommes sur les problèmes verbaux, alors que les participantes ayant résolu des problèmes numériques ont aussi bien réussi que les hommes. Pourtant, tous les problèmes numériques de cette étude étaient des conversions des problèmes verbaux utilisés. La difficulté des exercices proposés est extrêmement importante lorsque l'on évalue les effets de menace du stéréotype. MUZZATTI et AGNOLI (2007) ont par exemple suggéré que les filles et les garçons de leur première étude en école primaire ont obtenu les mêmes résultats en situation de menace du stéréotype à cause de la facilité des exercices de mathématiques proposés. Dans cette thèse, nous avons collaboré avec une équipe d'enseignants en mathématiques au collège pour créer les exercices de mathématiques. Nous les avons ensuite pré-testé auprès d'une classe de 6e et de 5e pour évaluer leur niveau de difficulté.

Ne pas connaître les différents modérateurs de la menace du stéréotype peut engendrer des difficultés à comprendre les résultats des études, en particulier lorsque les études ne répliquent pas les résultats habituels de la menace du stéréotype. Nous présentons dans la prochaine section les problèmes de réplicabilité et leurs implications dans la recherche sur la menace du stéréotype.

2.3.3. Biais méthodologiques pouvant causer des problèmes de réplication

Quelques auteurs s'interrogent sur la difficulté à répliquer les effets de menace du stéréotype en condition quasi-ordinaire (e.g., directement dans les écoles). Cela soulève plusieurs questions : est-ce que la menace du stéréotype existe dans la "vraie" vie ou est-ce un phénomène de laboratoire ? Y-a-t-il des biais dans les publications montrant les effets de menace du stéréotype ? Certains chercheurs ne répliquant pas les effets de menace du stéréotype remettent en cause l'existence du phénomène (FLORE et al., 2018 ; GANLEY et al., 2013 ; PENNINGTON et al., 2019). La contradiction et le débat scientifiques sont inévitables, souhaitables, et nécessaires à l'avancée de la recherche. Face à des résultats contradictoires, il est de coutume de procéder à une analyse comparative et critique des protocoles expérimentaux respectifs (caractéristiques des échantillons, méthodologie utilisée, type d'analyses statistiques effectuées), afin de comprendre l'origine des différences de résultats. Or, force est de constater que les études qui ne répliquent pas les effets de menace du stéréotype sont la plupart du temps citées sans s'interroger sur la méthodologie utilisée (pour des exceptions voir DANAHER et CRANDALL, 2008 ; RÉGNER et al., 2014 ; C. M. STEELE et DAVIES, 2003). Or, à y regarder de plus près, ces études ne sont pas exemptes de biais. Dans cette section, nous prenons les études de GANLEY et al. (2013) comme exemple, afin d'illustrer concrètement quelques-uns des biais méthodologiques. Ceci n'a pas pour but de stigmatiser ces auteurs. Nous aurions pu illustrer ces biais avec d'autres articles.

GANLEY et al. (2013) ont réalisé trois études en situation de classe avec des enfants entre 9 et 14 ans et des adolescents de 17 et 18 ans, et ont échoué dans ces trois études à répliquer les effets de menace du stéréotype chez les filles en mathématiques. Les auteurs s'interrogent alors ouvertement sur la réalité du phénomène, et d'autres auteurs reprennent leurs conclusions pour affirmer que les effets de menace du stéréotype ne sont pas répliqués en situation habituelle de classe (FLORE & WICHERTS, 2015). Or, plusieurs biais méthodologiques sont facilement repérables dans les études de GANLEY et al. (2013), et permettent d'expliquer l'absence de réplication des effets de menace du stéréotype. Le Tableau 2.1 présente les principaux éléments méthodologiques des études de GANLEY et al. (2013) que nous allons discuter.

TABLEAU 2.1

Participants, inductions expérimentales, et résultats obtenus dans chacune des trois études de GANLEY *et al. (2013).*

Études	Participants	Manipulation de la menace du stéréotype		Résultats
		Condition "menace du stéréotype"	Condition "menace réduite"	
Étude 1	212 élèves (102 garçons, 110 filles) de 13-14 ans, bon niveau scolaire	Avant le test de mathématiques, les élèves visionnent une vidéo qui présente des preuves scientifiques basées sur l'imagerie cérébrale expliquant que les compétences en mathématiques sont fixes et que les filles sont moins compétentes en mathématiques que les garçons. Plus précisément, le vocabulaire utilisé par le scientifique de la vidéo est le suivant : "recent research shows that math intelligence levels among students do not change as students get older. Students are born with a certain amount of natural math ability which does not change." (...) "females have lower levels of this kind of brain activity than males. This makes sense because girls often get lower scores on standardized tests compared to boys. The test that you will take today is a very good measure of your natural math ability."	Avant le test de mathématiques, les élèves visionnent une vidéo qui présente des preuves scientifiques basées sur l'imagerie cérébrale expliquant que les compétences en mathématiques sont malléables et que les filles et les garçons ont des compétences équivalentes en mathématiques. Plus précisément, le vocabulaire utilisé par le scientifique de la vidéo est le suivant : "recent research shows that among students, such as yourselves, math ability levels can and do change as students get older. We are finding that math ability is not just something you are born with. Math ability grows with practice, just like exercise strengthens your muscles. Students are capable of learning and mastering new math concepts at any time in their lives." (...) "males and females have equal levels of this kind of brain activity. This makes sense because young men and women, like yourselves, score the same on standardized math tests."	filles < garçons. Score moyen de réussite au test : 65%.
Étude 2	224 élèves (105 garçons, 119 filles) de 12-13 ans, 177 élèves (82 garçons, 95 filles) de 13-14 ans, bon niveau scolaire	Avant de faire le test de mathématiques, les élèves lisaient les instructions suivantes : "Please be sure to put an 'X' on the line next to your gender on the cover of the test booklet. This is very important, as boys have done much better than girls on this test in the past."	Les élèves recevaient uniquement les instructions au sujet du test de mathématiques.	Aucun effet significatif. Scores moyens de réussite très faibles pour tous les élèves : 24% pour les 12-13 ans et 34% pour les 13-14 ans.
Étude 3	68 élèves (39 garçons, 29 filles) de 9-10 ans, 105 élèves (40 garçons, 65 filles) de 13-14 ans, 145 élèves (69 garçons, 76 filles) de 17-18 ans	Les exercices du test de mathématiques étaient formulés de telle sorte à rendre le stéréotype de genre en mathématiques explicite : "At the Miller Middle School, the boys were much better at math than the girls. The math teachers chose the 20 students with the highest math test scores for the math team to represent the school at the statewide math competition. Eighteen of the students were boys and two were girls. What proportion of the students on the math team were boys?"	Les exercices du test de mathématiques étaient formulés de façon neutre, sans référence au stéréotype : "At the Miller Middle School, students were invited to participate in a special field trip, but there were only 20 spots available. The teachers chose 18 students from Ms. Fletcher's homeroom and two other students from Ms. Johnson's homeroom. What proportion of the students going on the field trip were from Ms. Fletcher's homeroom?"	filles < garçons. Scores moyens de réussite : 66% pour les 9-10 ans, 61% pour les 13-14 ans, 51% pour les 17-18 ans.

GANLEY et al. (2013) ont apporté plusieurs interprétations possibles à l'absence d'effet de menace du stéréotype dans les trois études : (1) la menace du stéréotype pourrait se manifester sous certaines conditions, mais il est difficile selon eux de comprendre lesquelles puisqu'ils n'ont pas retrouvé les effets de menace du stéréotype. (2) Les élèves de leurs études pouvaient présenter des caractéristiques pouvant modérer la menace du stéréotype. (3) Peut-être que les situations de test de mathématiques en général activent la menace du stéréotype, peu importe si l'on manipule une condition pour induire la menace du stéréotype. Il est surprenant que les auteurs soulèvent la troisième possibilité en 2013, alors que C. M. STEELE et DAVIES (2003) ont précisé que c'est la situation de test de mathématiques qui active par elle-même un effet de menace du stéréotype. Les auteurs ont commis un biais méthodologique pour les études 2 et 3, en considérant à tort que la condition standard de passation du test est une condition contrôle, sans menace du stéréotype. Dans ces deux études, la condition appelée "menace réduite" est en réalité une condition de menace du stéréotype dans laquelle le stéréotype a été activé implicitement par le simple fait qu'il s'agisse d'un test de mathématiques, et que les auteurs n'ont introduit aucune consigne ou action pour réduire l'effet de menace du stéréotype. Les auteurs ont donc comparés deux conditions de menace du stéréotype : une condition où le stéréotype est rendu explicite par les consignes (étude 2) ou le libellé du test de mathématiques (étude 3), et une condition où le stéréotype est activé implicitement par le fait de passer un test de mathématiques. Le fait que les filles aient obtenus de moins bonnes performances que les garçons dans l'étude 3 conforte la possibilité d'un effet de menace du stéréotype dans les deux conditions expérimentales. L'absence d'un tel effet de genre dans l'étude 2 est probablement dû à un effet plancher en raison d'une difficulté excessive du test pour tous les élèves (score moyen de réussite inférieur à 34%, alors que les scores moyens sont compris entre 51 et 66% dans les deux autres études). Contrairement aux études 2 et 3, la première étude comprend bien une situation de menace du stéréotype (vidéo sur les compétences innées en mathématiques) et une condition destinée à réduire la menace (vidéo sur la malléabilité des compétences en mathématiques). De ce fait, la présence d'un effet du genre des élèves (les filles ont moins bien réussi que les garçons dans les deux conditions expérimentales) pourrait indiquer que la vidéo destinée à réduire les effets de menace du stéréotype n'a pas eu l'effet escompté. Ceci a permis de révéler qu'une méthode fonctionnant chez les adultes (e.g., vidéo explicative de l'intelligence malléable, ARONSON et al., 2002) peut ne pas marcher chez les enfants. On peut toutefois s'interroger sur ce que les enfants de cette étude (12-14 ans) ont pu comprendre de chacune des deux vidéos. Malheureusement, aucune information n'est apportée à ce sujet. Aussi, il est important de s'interroger sur l'éthique de la méthode qui, sous prétexte d'activer la menace du stéréotype, renforce l'idée d'une conception innée de l'intelligence ainsi qu'une infériorité d'un groupe par rapport à un autre, en renforçant le stéréotype et notamment les explications biologiques des différences entre les filles et les garçons en mathématiques.

La publication de GANLEY et al. (2013) nous a servi d'exemple pour illustrer quelques-uns des biais méthodologiques sur lesquels les auteurs de la menace du stéréotype

avaient déjà attiré l'attention de la communauté scientifique dès 2003 (C. M. STEELE et DAVIES, 2003 ; voir aussi C. M. STEELE et ARONSON, 2004), et confortés par d'autres auteurs (CAMPBELL & COLLAER, 2009 ; SMITH & WHITE, 2002). D'autres biais méthodologiques existent, comme désamorcer l'enjeu évaluatif de la situation, par exemple car on ne veut pas stresser les élèves et impacter leurs notes (ARNOLD et al., 2020), ou encore en laissant un élève faire autre chose dans la salle de classe où l'expérimentation a lieu à cause de refus parental (FLORE et al., 2018).

Comme le soulignent RÉGNER et al. (2014), ne pas mentionner les biais méthodologiques des études susmentionnées peut conduire les chercheurs comme les enseignants à penser à tort que la menace du stéréotype n'existe pas en situation habituelle de classe ou d'examen (SACKETT et al., 2008). De plus, les résultats similaires des filles et des garçons sur les moyennes trimestrielles en classe viennent consolider, à tort, ces croyances. Comme bien d'autres effets, ceux liés à la menace du stéréotype sont en réalité plus faciles à observer en laboratoire qu'en situation habituelle de classe où les chercheurs et expérimentateurs ne peuvent pas contrôler nombre de variables parasites (e.g., un enseignant voulant rassurer sa classe en précisant que le test de mathématiques ne sera pas noté, retirant ainsi le stress de l'évaluation des capacités mathématiques des filles). Dans la mesure où les effets de menace du stéréotype sont très dépendants du contexte d'évaluation et de la nature des informations réactivées en mémoire, tester les effets de menace du stéréotype en situation réelle de classe représente une prise de risque. Les conditions à respecter pour faciliter l'observation des effets de menace du stéréotype en situation de classe quasi-ordinaire, mais également en laboratoire, sont les suivantes :

- Test de mathématiques complexes, mais pas trop pour ne pas avoir d'effet plancher.
- Situation évaluative à fort enjeu en s'assurant qu'aucun élément ne vienne neutraliser cette pression évaluative.
- Comparaison d'au moins deux conditions : menace du stéréotype *versus* menace réduite, en s'assurant que la deuxième condition contient bien une intervention pour réduire les effets des stéréotypes (ce n'est pas la condition standard de passation des tests).

2.3.4. Méthodes pour réduire les effets de menace du stéréotype

Au-delà de la démonstration de l'effet lui-même, les recherches sur la menace du stéréotype se sont très rapidement intéressées aux stratégies possibles pour en réduire les conséquences négatives sur les performances des groupes stigmatisés. L'enjeu, au niveau appliqué, est donc d'identifier des conditions de passation des tests qui puissent permettre à chacun et chacune de produire des performances à hauteur de ses compétences. Puisque le simple fait de passer un test de mathématiques important induit le phénomène de menace du stéréotype, sans qu'une mention explicite du stéréotype ne soit nécessaire, il faut trouver les moyens d'en réduire les effets. Depuis

les premières études de SPENCER et al. (1999) sur l'effet de menace du stéréotype chez les femmes en mathématiques, plusieurs stratégies ont été testées, que LIU et al. (2021) et SPENCER et al. (2016) ont classé en trois catégories : réduction de la perception de menace, stratégies de *coping*, et création d'un environnement sécurisant pour l'identité.

2.3.4.1. Réduction de la perception de menace

Cette catégorie d'interventions consiste à amener les individus du groupe stigmatisé (ici les filles et les femmes) à réévaluer la situation (évaluation en mathématique) comme étant moins menaçante qu'elle ne l'est en réalité. La consigne de falsification (i.e., présenter un test de mathématiques comme n'ayant jamais montré de différence de performance entre les hommes et les femmes) et la manipulation de la diagnosticité du test (i.e., présenter le test comme évaluant une autre compétence que celle ciblée par le stéréotype) sont les deux stratégies les plus représentatives de cette catégorie. La consigne de falsification (également appelée "*gender-fair instruction*") s'est révélée être efficace à la fois en laboratoire (BONNOT & CROIZET, 2011 ; BOUCHER et al., 2012 ; JOHN-HENDERSON et al., 2015 ; SPENCER et al., 1999) et en situation quasi-ordinaire de classe (GOOD et al., 2008 ; KELLER & DAUENHEIMER, 2003).

En laboratoire, BRODISH et DEVINE (2009) ont réussi à créer une situation de menace réduite en insistant sur le fait que le test de mathématiques de leur étude a été créé pour évaluer la mémoire de travail. Les femmes ont ainsi mieux réussi le test dans cette condition, comparé à celles ayant été notifié que le test mesurait leurs capacités naturelles en mathématiques. Quant à JOHN-HENDERSON et al. (2015), le test de mathématiques était décrit comme un test intellectuel évaluant les capacités en mathématiques dont les résultats entre hommes et femmes sont différents (situation menaçante) ou similaires (situation de menace réduite). Quelques participantes ont ensuite été encouragées à considérer l'anxiété comme utile lors des tests de performance. En situation menaçante, ces participantes ont obtenu de meilleurs résultats comparé à celles n'ayant reçu aucune information sur l'anxiété. Sans information supplémentaire sur l'anxiété, les auteurs ont observé les effets de menace du stéréotype habituels où les femmes en situation de menace ont obtenu de moins bons résultats que celles ayant lu l'instruction *gender-fair*. Cette différence de performance était beaucoup moins prononcée parmi les femmes ayant reçu les informations sur l'utilité de l'anxiété.

En situation de classe quasi-ordinaire, dans l'étude de GOOD et al. (2008), les étudiantes et les étudiants ont passé un test de mathématiques pendant leur cours, avec l'intervention d'un expérimentateur expliquant que le test mesurait les capacités en mathématiques. Pour la moitié des participants des deux groupes de genre, le test de mathématiques était précédé de la consigne de falsification. Les résultats ont confirmé l'effet bénéfique de cette consigne qui a permis aux femmes d'obtenir de meilleurs résultats que les étudiants (femmes et hommes) de la condition de menace du stéréo-

type (e.g., présentation du test de mathématiques sans la consigne de falsification). Chez les enfants, les études de HUGUET et REGNER (2007) et HUGUET et RÉGNER (2009) ont montré que les effets négatifs de menace du stéréotype étaient réduits lorsqu'un test a été présenté comme évaluant les capacités en dessin, contrairement à lorsqu'il a été présenté comme évaluant les capacités en géométrie. Les collégiens et collégiennes de 6e et de 5e devaient mémoriser la figure complexe de Rey, puis la dessiner sur papier. Les filles ont mieux retranscrit la figure que les garçons lorsque le test a été présenté comme un test de dessin. Dans cette situation de menace réduite, les mathématiques n'ont donc jamais été mentionnées.

Cependant, mentir sur la nature des compétences évaluées par un test ou affirmer l'absence de différences de performances entre hommes et femmes sont des pratiques difficilement utilisables en situation réelle de classe, d'examen, ou de concours. Il faudrait que les élèves et étudiants perçoivent un test de mathématiques comme n'étant pas des mathématiques, ou qu'ils pensent que l'on n'évalue pas leurs compétences en mathématiques. La phrase de falsification peut également échouer en situation réelle. Les effets de menace du stéréotype dépendent fortement du contexte situationnel. Il pourrait suffire qu'une personne exprime son désaccord à haute voix pour annuler les effets bénéfiques de cette consigne.

Une autre méthode consiste à présenter l'intelligence comme étant malléable. Le stéréotype de genre en mathématiques véhicule l'idée que les capacités mathématiques des filles et des femmes sont intrinsèquement inférieures aux garçons et aux hommes, ce qui sous-entend que les compétences différenciées en mathématiques selon le groupe de genre sont fixes. Considérer l'intelligence comme pouvant s'améliorer peut donc constituer un rempart important contre la menace du stéréotype. De plus, les étudiants pensant que l'intelligence ne peut changer sont plus susceptibles de se désidentifier d'un domaine après avoir reçu des retours négatifs sur leurs travails, parce qu'ils croient qu'ils ne pourront jamais changer (DWECK, 2000). Dans les deux études de GOOD et al. (2003), les élèves de 5e ont été mentorés par des étudiants à l'université. Le mentorat expliquant que l'intelligence est extensible a permis de supprimer les écarts de performance entre les filles et les garçons en mathématiques. En apprenant progressivement, les élèves ont découvert l'importance des efforts et de la motivation, tout en diminuant les croyances sur le talent naturel en mathématiques. De plus, ceux qui valorisent davantage les efforts que les capacités innées sont moins enclins à craindre la confirmation de stéréotypes négatifs. D'autres recherches indiquent que concevoir l'intelligence comme malléable et non innée profite à tous les élèves (BLACKWELL et al., 2007). Cette perspective aide également à faire comprendre aux élèves que la performance ne reflète pas nécessairement la compétence, et les encourage à relativiser en cas d'échec pour éviter de se désidentifier des mathématiques.

2.3.4.2. Stratégies de *coping*

Comme mentionné par SPENCER et al. (2016), cette catégorie de stratégies ne permet pas de réduire le niveau de menace de la situation de test, mais permet en revanche d'aider les femmes et les filles à mieux gérer la menace pour que leurs performances ne soient pas affectées. L'une de ces méthodes, appelée "auto-affirmation", consiste à encourager les femmes à réfléchir à leurs qualités et à les lister afin de renforcer leur confiance en elles. Ces qualités peuvent même ne pas être en rapport avec les mathématiques ; il s'agit ici de réaffirmer leur image de soi en général. Ainsi, les participantes de l'étude de MARTENS et al. (2006) ayant classé leurs qualités tout en écrivant leur importance pour elles-mêmes ont obtenu les mêmes résultats que les hommes participants. Cela n'a eu aucun impact sur les performances des hommes. Les chercheurs ont suggéré que l'auto-affirmation est efficace pour réduire les effets de la menace générée par le stéréotype négatif, mais pas forcément pour améliorer les performances de n'importe quel individu. Écrire ses qualités est en fait une méthode pour empêcher les pensées négatives d'interférer avec la tâche en cours. Cela réduit donc la possibilité de la menace du stéréotype d'agir sur la capacité de la mémoire de travail. LOGEL, ISERMAN et al. (2009) ont fourni des indications aux participants pour éviter les pensées négatives, en leur demandant d'écrire quelque chose d'important à propos de leur identité. Ceci a permis aux femmes d'obtenir les mêmes résultats que les hommes. Les femmes n'ayant pas eu d'instructions pour supprimer les pensées négatives ont moins bien réussi que les hommes. Par ailleurs, BEILOCK et al. (2007) ont montré dans leurs 5 études que la menace du stéréotype agit lorsque les exercices de mathématiques dépendent fortement des ressources de la mémoire de travail. Ils ont alors montré que la menace du stéréotype peut être atténuée en pratiquant intensivement ces exercices de mathématiques, de sorte à utiliser la mémoire à long terme plutôt que la mémoire de travail.

Une autre méthode consiste à mettre en valeur une identité sociale chez les femmes/filles qui ne serait pas ciblée par un stéréotype négatif dans le domaine des mathématiques. SHIH et al. (1999) ont évalué si le stéréotype portant sur les dons naturels en mathématiques des personnes Asiatiques pouvait réduire les effets de menace du stéréotype chez les femmes d'origine Asiatique. En activant subtilement une de leur identité sociale (e.g., en leur demandant de préciser leur origine pour activer leur identité d'origine culturelle, et en leur demandant de préciser leur genre pour activer leur identité de genre) avant de passer un test de mathématiques, les auteurs ont montré que les participantes ont mieux réussi lorsque leur identité d'origine culturelle a été mise en avant. Ces résultats ont été répliqués chez les enfants de 7 à 14 ans par AMBADY et al. (2001). Ces études ont testé l'appartenance à un autre groupe social ciblé par un stéréotype, montrant encore une fois l'impact des stéréotypes sociaux sur la performance en mathématiques. RYDELL et al. (2009) ont quant à eux mis en valeur l'appartenance des femmes aux étudiants à l'université. Dans leur étude, toutes les participantes étaient étudiantes à l'université et ont passé un test de mathématiques décrit comme un moyen d'examiner pourquoi les femmes sont généralement moins

douées en mathématiques. En situation de menace réduite, il a été ajouté que le test est également un moyen de comprendre pourquoi les étudiants à l'université sont meilleurs en mathématiques que les personnes n'étant pas à l'université. La mise en évidence de l'appartenance des participantes à un groupe ayant des avantages en mathématiques a ainsi permis de réduire les effets de menace du stéréotype. De même, McGlone et Aronson (2006) ont montré que lorsque l'appartenance aux étudiants en école privée a été mise en avant, les femmes ont bien mieux réussi le test de mathématiques comparé à celles dont l'identité de genre a été mise en avant. Il est intéressant de noter que c'était le contraire pour les hommes, bénéficiant du boost du stéréotype.

Sans mettre en avant une appartenance à un groupe, une autre méthode consiste plutôt à estomper les stéréotypes de genre en brouillant les possibles différences entre les hommes et les femmes. Contrairement à la consigne de falsification qui précise que les hommes et les femmes obtiennent les mêmes résultats à un test de mathématiques, il s'agit ici de ne pas parler des performances en mathématiques et de mettre plutôt en avant les points communs entre ces deux groupes de genre (i.e., les appartenances croisées). Le fait de lister les caractéristiques communes entre les hommes et les femmes avant de passer un test de mathématiques permet de réduire les effets de menace du stéréotype (Rosenthal & Crisp, 2006; Rosenthal et al., 2007). En allant plus loin que "flouter" son genre, J.-E. R. Lee et al. (2014) et Ratan et Sah (2015) ont constaté qu'incarner un avatar masculin lors de test de mathématiques compétitifs a permis aux participants, femmes comme hommes, d'avoir de meilleurs résultats que ceux incarnant un avatar féminin. Les auteurs ont ainsi attribué un autre groupe social aux femmes (e.g., en incarnant un homme). Sous l'influence de sa propre image virtuelle, les individus tendent en effet à auto-influencer leurs attitudes et comportements dans un environnement virtuel, en s'alignant avec les caractéristiques identitaires de son avatar, un phénomène appelé effet Proteus (N. Yee & Bailenson, 2007). Les femmes des études de J.-E. R. Lee et al. (2014) et Ratan et Sah (2015) ont sûrement bénéficié du boost du stéréotype, un phénomène se produisant en situation de menace du stéréotype où les résultats des hommes augmentent, car le stéréotype est en leur faveur (H. J. Johnson et al., 2012; Walton & Cohen, 2003). Cette méthode est intéressante pour montrer comment les stéréotypes impactent les performances en mathématiques, mais n'est pas réalisable en situation réelle car cela risque de renforcer ces stéréotypes, où les individus peuvent penser que seuls les hommes sont capables de réussir en mathématiques.

Bien que renier le stéréotype négatif sur les performances en mathématiques des femmes ne suffit pas à écarter la menace du stéréotype, connaître ce phénomène peut permettre aux femmes d'en réduire ses effets, comme montré par Johns et al. (2005). Dans leur étude, les participantes ayant été informées que l'anxiété qu'elles pourraient ressentir lors du test de mathématiques était due aux stéréotypes négatifs répandus dans la société, et que cela n'avait aucun rapport avec leur capacité à réussir le test, ont mieux réussi que les femmes qui pensaient participer à une étude sur les différences

de performances en mathématiques entre hommes et femmes. Elles ont ainsi obtenu les mêmes résultats que les hommes. Cette stratégie consiste en fait à attribuer la faute d'une possible contre-performance à un facteur externe. Dans la deuxième étude de BEN-ZEEV, FEIN et INZLICHT (2005), les participantes entendaient un léger bruit lors d'un test de mathématiques. Pour certaines, l'expérimentateur a précisé que le bruit subliminal augmentera leur nervosité et leur fréquence cardiaque, permettant ainsi d'attribuer leur stress à ce bruit. Pour d'autres, il a été précisé que le bruit subliminal n'aura aucun impact sur elle. Les participantes pensant que le bruit augmenterait leur fréquence cardiaque ont toutes obtenu des résultats similaires, qu'elles étaient en situation de menace du stéréotype (en présence de deux hommes) ou en situation de menace réduite (en présence de deux femmes). En revanche, lorsqu'elles n'ont pas eu la possibilité d'attribuer leur stress au bruit, les femmes en situation de menace du stéréotype ont moins bien réussi que celles en situation de menace réduite.

SMEDING et al. (2013) ont testé une méthode ne nécessitant aucune intervention de la part des élèves, en modifiant l'ordre de passation d'un contrôle de français et de mathématiques dans une même journée scolaire dans plusieurs collèges en France. Les filles ayant passé le test de français avant de passer le test de mathématiques ont aussi bien réussi que les garçons en mathématiques, alors qu'elles ont eu des résultats inférieurs aux garçons au test de mathématiques lorsque l'ordre était inversé. De plus, l'ordre des tests n'a pas affecté la performance des filles au test de français, ni les performances des garçons dans les deux tests. Autrement dit, passer un test de français avant un test de mathématiques permet aux filles d'avoir une meilleure note en mathématiques, sans que cela impacte les garçons. Les filles étant négativement stéréotypées en mathématiques, mais positivement stéréotypées en français par rapport aux garçons (SHIBLEY HYDE & KLING, 2001), elles ont eu la possibilité d'affirmer leur image et leur confiance en elles dans un autre domaine que les mathématiques (MARTENS et al., 2006). Comme nous l'avons vu précédemment, la mise à disposition d'informations positives aux filles et aux femmes les aident à contrer l'impact du déséquilibre entre leur besoin de renforcer leur estime de soi et les stéréotypes suggérant que les filles/femmes sont moins performantes en mathématiques et devraient échouer dans ce domaine (MCINTYRE et al., 2003 ; RYDELL et al., 2009 ; SMEDING et al., 2013 ; TAILLANDIER-SCHMITT et al., 2012).

Ces méthodes peuvent être très utiles pour les filles et les femmes en situation réelle, ainsi que pour les enseignants qui pourraient leur apprendre ces stratégies. Cependant, cela ne change pas l'environnement de travail où la menace du stéréotype reste présente, et où l'identité de genre des filles et des femmes reste menacée. Nous présentons dans la prochaine sous-section des études permettant de créer des environnements plus sécurisants pour réduire les effets de menace du stéréotype.

2.3.4.3. Création d'un environnement sain pour protéger l'identité de genre : les modèles de réussite

Les méthodes que nous présentons dans cette section changent l'environnement pour réduire les effets de menace du stéréotype, en créant un environnement plus serein. Si l'environnement n'est pas hostile (e.g., grande majorité de filles présentes lors du test, HUGUET et REGNER, 2007 ; PICHO et STEPHENS, 2012), la menace du stéréotype agit moins sur les filles. Au contraire, s'il est hostile (e.g., grande majorité de garçons présents lors du test, J.-E. R. LEE et NASS, 2012, affiche dégradante sur la capacité des filles en mathématiques, OSWALD et HARVEY, 2000), les effets négatifs de menace du stéréotype sont amplifiés. Une façon de créer un environnement plus serein pour les filles et les femmes en mathématiques est de mettre en valeur leur identité de genre, en montrant que celle-ci n'est pas un obstacle à la réussite en mathématiques. Nous avons précédemment montré quelques études mettant en valeur d'autres identités que le genre (e.g., catégorie sociale, origine culturelle) pour augmenter les performances des filles en mathématiques. Il s'agit ici de réaffirmer les qualités des femmes de manière générale en présentant un **modèle de réussite**. Des études ont montré qu'un modèle de réussite, compétent en mathématiques, peut réduire les effets négatifs de menace du stéréotype en s'identifiant à ce modèle. Dans cette thèse, nous nous concentrons sur les modèles de réussite pour induire une situation de menace réduite. Par exemple, dans l'étude de LUONG et KNOBLOCH-WESTERWICK (2017), les participantes ont pu parcourir des magazines montrant des modèles sociaux féminins dans des domaines stéréotypés (beauté ou famille) ou contre-stéréotypés (carrière ou sciences) avant de passer un test de mathématiques, en situation de menace du stéréotype (où il a été spécifié que le test mesurait les compétences en mathématiques) ou de menace réduite (où il a été spécifié que le test mesurait les compétences de raisonnement). Les résultats ont montré que les femmes passaient plus de temps sur les magazines contre-stéréotypés lorsqu'elles étaient en situation de menace du stéréotype, et l'effet des modèles sur leurs performances était modéré par l'assimilation des participantes aux modèles : plus elles s'identifiaient aux modèles, plus leurs notes étaient élevées.

En effet, lorsque les individus ressentent de l'anxiété, par exemple face à une tâche perçue comme trop difficile, une stratégie consiste à chercher à s'affilier à d'autres personnes (SCHACHTER, 1959). L'objectif est d'établir des liens en percevant des similitudes, objectives ou non (SELFHOUT et al., 2009), entre les caractéristiques de l'autre et les nôtres (TESSER, 1986). Cette personne à laquelle nous nous affiliions devient alors un modèle social. Les attributs recherchés chez ces modèles sociaux incluent leur statut et leurs compétences. C'est pourquoi les apprenants tendent à se rattacher à des individus un peu plus âgés ayant un statut plus élevé, tout en partageant des caractéristiques similaires telles que le genre ou l'origine culturelle (BANDURA, 1986 ; MUSSWEILER, 2003). Autrement dit, l'apprenant aspire à s'inspirer de cette personne et à lui ressembler. L'influence des modèles de réussite peut cependant être négative si la réussite de ces modèles est perçue comme inatteignable (HOYT & SIMON, 2011),

ou encore si la situation professionnelle du modèle semble peu attractive ou moins facilement atteignable que d'autres (BREDA et al., 2019). MARX et KO (2012) ont trouvé que lorsqu'un modèle de réussite féminin était très compétent en mathématiques, plus il y avait de similarités (e.g., provenir de la même école, avoir les mêmes centres d'intérêts) entre les participantes et le modèle de réussite, moins ils étaient perçus comme intimidants, et plus les performances des participantes au test de mathématiques étaient élevées.

Dans la deuxième étude de MCINTYRE et al. (2003), les participants ont passé un test de mathématiques où il a été précisé que des études ont montré que les hommes obtiennent de bien meilleurs résultats que les femmes en mathématiques. Il a également été précisé que quelques études ne répliquent pas ce résultat. Une partie des participants ont alors lu des vraies histoires de quatre femmes réussissant en médecine, droit, ingénierie et biologie. Le reste des participants ont lu des histoires de quatre entreprises ayant réussi. Les femmes ayant lu les histoires sur les entreprises ont obtenu les moins bonnes notes. Les hommes et les femmes ont obtenu des résultats similaires lorsqu'ils ont lu les histoires des modèles de réussite féminin. Notamment, les femmes ayant lu ces histoires ont mieux réussi que les autres femmes ayant lu les histoires sur les entreprises. Les performances des femmes face à la menace du stéréotype en mathématiques peuvent être ainsi améliorées en leur rappelant le succès d'autres femmes, sans que cela n'ait d'incidence sur les performances des hommes. MARX (2019) a quant à lui plus particulièrement étudié l'influence des pairs sur la performance en mathématiques des étudiantes. Des précédentes recherches ont en effet montré que les étudiantes sont plus sensibles à l'influence de leurs pairs, notamment en suivant des cours de mathématiques lorsque leurs amis proches (filles ou garçons) ou leurs camarades de classe y réussissent (CROSNOE et al., 2008). Lors des deux études de MARX (2019), les participantes ont lu un article de journal fictif sur un homme ou une femme à l'université, tous compétents en mathématiques, et provenant de l'école des participants ou provenant d'une autre école. Elles ont ensuite effectué un test de mathématiques dont il a été précisé que les hommes et les femmes produisaient des résultats différents. Les deux études montrent que lorsque le pair masculin provient d'une école différente, il est suffisamment éloigné des participantes et cesse d'être perçu comme menaçant. Il peut alors améliorer les performances des filles au même titre que les pairs féminins. Cette étude suggère que les modèles masculins peuvent également améliorer la performance des filles, mais sous certaines conditions. D'autres études ont montré que les modèles de réussite masculins baissent les performances des filles/femmes car perçus comme menaçants, alors que les modèles de réussite féminins perçus comme très compétents plutôt que modérément compétents baissent les performances des hommes car ces modèles remettent le stéréotype en question (MARX & ROMAN, 2002). De plus, dans l'étude de MARX et al. (2013), lorsqu'un modèle social de même genre que le participant/la participante exprimait des inquiétudes quant à la façon dont on peut le percevoir (e.g., « durant ma scolarité, j'ai eu peur que si j'échouais, les gens penseraient que c'est parce qu'une personne comme moi ne peut pas réussir en mathématiques »),

cela augmentait la performance des hommes, et baissait celle des femmes. Chez les hommes, le modèle féminin exprimant des inquiétudes leur a permis d'avoir de meilleurs résultats que le modèle masculin n'exprimant aucune inquiétude. Chez les femmes, le modèle féminin n'exprimant aucune inquiétude leur a permis d'avoir de meilleures performances que tous les autres modèles (masculins et féminin exprimant des inquiétudes). Un modèle féminin exprimant des doutes quant à sa réussite ne va pas à l'encontre du stéréotype négatif, ce qui pourrait expliquer pourquoi ces modèles ont été moins efficace sur la performance en mathématiques des femmes. Ces doutes et inquiétudes n'ont pas besoin d'être exprimés oralement pour avoir un effet négatif. En effet, dans l'étude de BEILOCK et al. (2010) sur une année en école primaire, lorsque le comportement non-verbal des femmes enseignantes en mathématiques exprimait de l'anxiété, cela réduisait les performances en mathématiques et les croyances stéréotypiques des filles.

BAGÈS et MARTINOT (2011) ont étudié les différents discours des modèles de réussite permettant d'améliorer les performances des participantes d'une dizaine d'années en situation de menace du stéréotype. Plus précisément, les performances à un test de mathématiques s'amélioraient lorsque les jeunes apprenants lisaient une histoire d'un modèle de réussite en mathématiques, juste avant le test. Le modèle de réussite était un garçon ou une fille, compétent en mathématiques et légèrement plus âgé que l'élève cible, et la cause de sa réussite en mathématiques a été expliquée soit par son travail, soit par ses capacités naturelles, soit aucune explication n'était fournie. Les performances des garçons comme pour les filles étaient meilleures avec le modèle de réussite féminin. En outre, lorsque le modèle expliquait qu'il a eu de bons résultats en mathématiques grâce à son travail et ses efforts, les performances des apprenants étaient encore meilleures, indépendamment du genre du modèle et de l'apprenant. Ces résultats sont cohérents avec l'étude de MARX (2019) : les modèles masculins perçus comme étant moins menaçants permettent également d'améliorer les performances des filles et des femmes. De plus, les résultats de BAGÈS et MARTINOT (2011) sont cohérents avec les études montrant les bienfaits de préciser que l'intelligence n'est pas innée, et que les capacités peuvent croître avec des efforts et du travail (études présentées en Section 2.3.4.1). Dans une étude similaire, les résultats de BAGÈS et al. (2016) ont également montré que les filles ont aussi bien réussi que les garçons avec un modèle de réussite, peu importe son genre, expliquant qu'il a travaillé pour réussir en mathématiques. Les filles ont moins bien réussi que les garçons avec le modèle de réussite naturellement doué. De plus, les filles comme les garçons se sont plus identifiés aux modèles ayant travaillé pour réussir, en comparaison au modèle naturellement doué et au modèle n'expliquant pas comment il a pu réussir en mathématiques.

Les filles et les femmes doivent en effet pouvoir distinguer que le modèle a mérité de réussir, notamment par son travail, pour contrecarrer la croyance en des capacités mathématiques innées (C. M. STEELE, 1997), et par conséquent, contrecarrer les effets de menace du stéréotype. Dans la première étude de MCINTYRE et al. (2011), les parti-

cipantes ont ainsi mieux réussi après avoir lu l'histoire d'une femme ayant réussi par ses propres compétences et son travail acharné, et ayant eu l'idée d'inventer des chauffages respectueux de l'environnement. Les participantes ayant lu une histoire sur une femme dont son défunt époux était le véritable inventeur du chauffage respectueux de l'environnement, et qui lui a fait promettre à contrecœur de déposer les brevets sous le nom de sa femme, ont obtenu de moins bons résultats que les premières participantes. Elles ont également obtenu les mêmes résultats que les participantes n'ayant pas lu d'histoire. Dans leur deuxième étude, plus les modèles de réussite étaient méritants en montrant un succès stable et constant, plus les performances des participantes étaient élevées. Des résultats similaires ont été trouvé par TAYLOR et al. (2011), où les femmes percevant Hillary Clinton (le modèle social présenté dans l'étude) comme méritante ont obtenu de meilleurs résultats que les femmes qui la percevaient comme non-méritante.

Outre le genre, le discours et les gestes, d'autres caractéristiques comme la féminité peuvent entraver l'efficacité des modèles de réussite. Les deux études de BETZ et SEKAQUAPTEWA (2012) ont montré que, parmi les filles d'une dizaine d'années étant peu identifiées aux mathématiques, la féminité perçue d'une femme compétente en STIM, portant du rose et aimant lire des magazines de mode, réduisait leur intérêt en mathématiques, leur sentiment de pouvoir réussir dans une future carrière en sciences, et leur sentiment de réussite, comparé à une même femme ayant une apparence et des centres d'intérêt perçus comme plus neutres (e.g., porter des vêtements foncés, aimer lire). Ces résultats se retrouvent chez les jeunes filles fortement identifiées aux mathématiques, à l'exception de leur sentiment de pouvoir réussir une carrière en STIM, où le modèle très féminin et compétent en mathématiques n'impacte pas négativement ce sentiment. En revanche, ces résultats ne sont pas répliqués lorsque le modèle de réussite n'est pas seulement compétent en STIM, mais est compétent à l'école. Ceci indique que ce ne sont pas les signes de féminité qui entraînent des effets négatifs, mais une combinaison de la féminité et du succès en STIM qui semble inatteignable pour les jeunes filles.

En effet, même si les modèles contre-stéréotypiques peuvent inspirer, ils n'affaiblissent pas nécessairement les stéréotypes de genre. Par exemple, les modèles très féminins peuvent être considérés comme des exceptions à la règle (Z. RICHARDS & HEWSTONE, 2001). Ils peuvent également être recatégorisés dans un sous-ensemble de leur groupe social pour les écarter, et ainsi maintenir les stéréotypes (KUNDA & OLESON, 1995). Les femmes qui excellent dans des domaines stéréotypés masculins comme les STIMs sont ainsi classées dans une catégorie de genre moins féminine (HEILMAN & OKIMOTO, 2007). Par exemple, dans l'étude de KESSELS (2005), des adolescents de 15 ans ont attribué moins de traits féminins ainsi que plus de traits masculins aux étudiants aimant la physique (un domaine perçu comme masculin), en comparaison aux étudiants aimant la musique (un domaine perçu comme féminin). Le stéréotype selon lequel les femmes réussissant en STIM ne seraient pas féminines s'appliquent également à l'apparence, où le maquillage est perçu comme incompa-

tible avec la réussite en mathématiques (PRONIN et al., 2004). Cependant, dans les études de CHERYAN, SIY et al. (2011), le sentiment de réussite des femmes envers les STIMs a tout de même baissé lorsqu'elles ont interagi avec une femme non-féminine représentant un modèle de réussite en STIM. Cela s'explique par les goûts et l'apparence geek de ce modèle : les participantes ne se sont pas identifiées à elle, jugée comme trop stéréotypée par rapport aux sciences, et trop éloignée d'elles-mêmes. Ainsi, même une femme masculine représentant un modèle de réussite peut ne pas être efficace pour lutter contre la menace du stéréotype, bien qu'aucune de ces études n'aient testé l'impact des femmes féminines ou masculines en tant que modèle de réussite sur la performance en mathématiques des filles.

En résumé, un modèle de réussite doit être un individu compétent dans son domaine auquel l'élève peut s'identifier pour induire l'idée que lui aussi peut y arriver. Des caractéristiques importantes sont cependant à prendre en compte pour éviter une contre-performance en mathématiques. Par exemple, le modèle de réussite doit être légèrement plus âgé que l'élève. En effet, si le modèle de réussite est plus jeune ou a le même âge que les élèves, ceux-ci peuvent se sentir inférieurs à lui s'ils n'ont pas les mêmes résultats en mathématiques que le modèle. Par ailleurs, si le modèle est trop âgé, l'enfant ne pourra pas s'identifier (BAGÈS et al., 2016) car le modèle pourrait être perçu comme n'appartenant pas au même groupe social (J. STEELE, 2003). Le genre joue également un rôle important dans la définition de modèle de réussite chez les enfants : les garçons s'identifient aux garçons, les filles s'identifient aux filles. Cependant, d'après les études de BAGÈS et MARTINOT (2011), expliquer que le modèle de réussite est parvenu à avoir de bons résultats en mathématiques grâce aux efforts et au travail, peu importe son genre, permet aux élèves d'augmenter leurs performances en mathématiques. Par ailleurs, LESLIE et al. (2015) ont montré qu'il y a des inégalités de genre dans les domaines où les personnes croient qu'avoir un don/un talent inné est un élément clé pour réussir, ce qui est le cas notamment en mathématiques. Les étudiants de cette étude évaluaient les femmes comme pouvant moins bien réussir en sciences que les hommes car d'après eux, elles ne sont pas naturellement douées pour cela. Contrer cette croyance permet également de contrer la menace du stéréotype. Aussi, bien que des résultats positifs en mathématiques chez les femmes peuvent résulter d'une seule exposition à un modèle féminin, ils peuvent également se développer progressivement à mesure que les élèves continuent d'interagir avec le modèle (STOUT et al., 2011). Des études ont également montré que les filles perçoivent plus facilement d'autres filles en tant que modèle de réussite après avoir établi des connections personnelles avec elles (BUCK et al., 2008 ; DASGUPTA & ASGARI, 2004). Plus précisément, LEROY et al. (2022) ont réalisé une étude longitudinale directement dans les classes où l'élève ayant les meilleurs résultats en mathématiques, faisant office de modèle de réussite dont l'exposition est continue, était une fille ou un garçon. Lorsque les filles sont dans une classe où une fille avait les meilleurs résultats en mathématiques tout au long de l'année scolaire, elles obtenaient de meilleurs résultats à la fin de l'année, contrairement aux filles dans une classe où un garçon avait les meilleures notes en mathématiques. Cette étude montre l'importance d'une

exposition continue à un modèle de réussite ayant d'une part des similitudes avec les filles, et pouvant d'autre part interagir fréquemment avec elles.

2.4. Discussion

Les stéréotypes de genre influencent négativement la performance des filles et des femmes en mathématiques à travers le phénomène appelé menace du stéréotype. Plusieurs interventions ont été testées par les chercheurs pour en réduire les effets lors de tests de mathématiques (Section 2.3.4). Parmi ces méthodes, l'exposition aux modèles de réussite féminins compétents en mathématiques permet aux filles et aux femmes de réaffirmer leur image, à condition que la réussite de ces modèles soit perçue comme atteignable. Le domaine des mathématiques étant sur-représenté par des pairs masculins, cela renforce davantage le sentiment de non-appartenance que les femmes peuvent ressentir en sciences. Les pairs féminins peuvent réduire les effets de menace du stéréotype et le sentiment de non-appartenance en sciences que les autres femmes peuvent ressentir, lorsqu'elles sont perçues comme des modèles de réussite méritantes auxquels les filles et les femmes peuvent s'identifier (Section 2.3.4.3). Réduire les effets de menace du stéréotype à l'aide de modèle de réussite est particulièrement intéressant dans le domaine de l'éducation. Cette méthode permet d'augmenter la performance en mathématiques des filles, sans pour autant baisser celle des garçons. Certaines méthodes, comme spécifier que les filles et les garçons obtiennent le même score dans un test de mathématiques, peuvent en effet parfois réduire la performance des garçons et des hommes, se sentant moins en confiance que d'habitude car le stéréotype qui les favorise est mise à mal.

Une limite que l'on retrouve dans la plupart des études sur l'impact des modèles de réussite sur la performance en mathématiques des filles et des femmes est que les participantes ne pouvaient interagir avec les modèles, puisqu'il s'agissait le plus souvent de lire leurs histoires avant de passer un test de mathématiques. Or, une interaction continue avec un modèle de réussite permet de développer progressivement les performances des élèves. De plus, des études ont montré que les filles percevaient des femmes scientifiques en tant que modèle de réussite après avoir établi des connections personnelles avec elles. Il faut cependant faire attention aux comportements verbaux et non-verbaux du modèle de réussite, car ils peuvent produire des effets négatifs sur l'apprentissage, par exemple s'ils véhiculent de l'anxiété ou des doutes par rapport au niveau de compétence en mathématiques. Aujourd'hui, les modèles féminins de réussite en sciences sont mis en avant dans la société, par exemple par Femmes et Sciences avec les portraits scientifiques "La Science taille XX-Elles", ou encore la fondation l'Oréal avec son programme de sensibilisation "For Girls in Science". La mise en place de ce genre d'actions nécessitent de bien prendre en compte différents facteurs d'influences. Par exemple, des analyses sur l'efficacité du programme de l'Oréal ont montré que l'influence sur les orientations des jeunes filles ne repose pas exclusivement sur la capacité des modèles de réussite féminins à

démystifier les stéréotypes de genre liés aux métiers scientifiques, mais également sur l'attrait de leur situation professionnelle pour les jeunes filles. Les modèles de réussite ayant une situation professionnelle perçue comme plus attractive et plus facilement atteignable par les lycéennes ont facilité l'identification à ces modèles, qui a influencé les choix d'orientation des jeunes filles vers une classe préparatoire scientifique. Autrement dit, l'impact du programme sur les choix d'orientation des filles dépend du profil des modèles de réussite féminins qui ont conduit les interventions en classe.

En outre, l'exposition -continue ou non- à un modèle de réussite pour réduire les effets de menace du stéréotype est compliqué à mettre en place en situation de classe réelle. Il faudrait que des modèles de réussite interviennent en personne dans chaque classe juste avant un test de mathématiques, en ayant des similitudes avec les élèves, et en étant attentifs à leur propre comportement verbal et non-verbal pour ne pas produire l'effet inverse. De plus, étant donné le nombre limité de femmes en mathématiques, certaines études ont préféré analyser le moment le plus opportun pour exposer les filles et les femmes à un modèle de réussite féminin. Puisque la menace du stéréotype influence plus particulièrement les individus fortement identifiés au domaine, il serait plus judicieux de présenter ces modèles à ces élèves pour les retenir dans le domaine des sciences pour être sûr d'utiliser le temps de ces modèles à bon escient. Cependant, ceci mettrait de côté les filles peu identifiées au domaine qui pourraient tout de même bénéficier des modèles de réussite féminins.

Le recours aux personnages virtuels incarnant des modèles de réussite permettrait de pallier ces difficultés. Le numérique étant de plus en plus utilisé, dans les classes comme à la maison, y inclure un personnage virtuel incarnant un modèle de réussite permettrait d'aider à la pédagogie des enseignants pour lutter contre la menace du stéréotype, tout en améliorant les performances des filles et en leur permettant d'affirmer leur image. Dans cette thèse, nous nous intéressons à la création et à l'évaluation de ce type de personnage virtuel. Le prochain chapitre est consacré aux personnages virtuels pédagogiques, et plus particulièrement à l'impact du genre, du rôle, et des interactions de ces personnages dans les domaines académiques.

2.5. Points clés du chapitre

- Il existe des différences de performances en mathématiques en faveur des garçons et des hommes lors des tests standardisés difficiles.
- Le stéréotype de genre véhiculé dans nos sociétés suppose que les filles et les femmes sont naturellement moins douées en mathématiques que les garçons et les hommes.
- La menace du stéréotype correspond à la crainte d'être négativement stéréotypé lors d'un test de compétence ou de confirmer un stéréotype négatif envers le groupe social auquel on appartient, ce qui peut conduire à des contre-performances.
- La menace du stéréotype explique en partie les différences de performances en mathématiques en faveur des garçons et des hommes, chez les enfants comme chez les adultes.
- Plusieurs techniques existent pour réduire les effets de menace du stéréotype, comme présenter des modèles de réussite incarnés par des filles ou des femmes compétentes en mathématiques.
- Un modèle de réussite correspond à une personne compétente dans son domaine, légèrement plus âgée, à laquelle on s'identifie à travers des similitudes (genre, origine culturelle, loisirs, ...).
- Quelques études ont montré que le genre du modèle de réussite n'a pas d'importance dès lors qu'il explique sa réussite par du travail et des efforts.
- Il est difficile de faire intervenir des vrais modèles de réussite tout le temps dans les classes.
- Notre objectif est de créer des personnages virtuels pouvant incarner des modèles de réussite pour tester leur efficacité contre les effets de menace du stéréotype chez les collégiennes de 6e et 5e.

3. Agents virtuels pédagogiques : modélisation du genre, du rôle de modèle de réussite, et du rapport virtuel

Sommaire

3.1. Introduction . . . 65
3.2. Impact du genre des agents virtuels sur la perception et la performance des utilisateurs . . . 68
3.2.1. Méthode de la revue systématique . . . 68
3.2.1.1. Stratégie de recherche d'articles . . . 68
3.2.1.2. Sélection des articles . . . 69
3.2.2. Évaluations des agents virtuels . . . 69
3.2.2.1. Évaluations de la perception des utilisateurs . . . 69
3.2.2.2. Évaluations objectives de l'apprentissage des utilisateurs 71
3.2.2.3. Évaluations subjectives de l'apprentissage des utilisateurs 71
3.2.3. Persistance des stéréotypes de genre dans les interactions humain-agent . . . 72
3.2.4. Effet du genre des agents virtuels pédagogiques sur les résultats scolaires . . . 91
3.3. Agents virtuels pédagogiques en tant que modèles de réussite . . . 102
3.4. Le rapport virtuel pour améliorer les performances académiques . . . 106
3.4.1. Fondement théorique du rapport . . . 106
3.4.1.1. La positivité . . . 109
3.4.1.2. L'attention mutuelle . . . 110
3.4.1.3. La coordination . . . 111
3.4.2. Modélisation du rapport virtuel . . . 112
3.5. Vers un modèle d'agent virtuel pédagogique pour lutter contre la menace du stéréotype . . . 118
3.6. Discussion . . . 121
3.7. Points clés du chapitre . . . 123

Ce chapitre est adapté de :

Armando, M., Ochs, M., & Régner, I. (2022). The Impact of Pedagogical Agents' Gender on Academic Learning : A Systematic Review. *Frontiers in Artificial Intelligence*, 5, 862997. https://doi.org/10.3389/frai.2022.862997 (ARMANDO et al., 2022)

Armando, M., Régner, I., & Ochs, M. (2021). Les effets du genre des agents virtuels sur l'apprentissage. *Workshop sur les "Affects, Compagnons Artificiels et Interactions" (WACAI)*, Île d'Oléron, France. https://hal.science/hal-03366193v2 (ARMANDO et al., 2021)

3.1. Introduction

Les *outils numériques pour l'éducation* pour aider à l'apprentissage sont de plus en plus utilisés, en classe comme à la maison, notamment depuis la pandémie de la Covid-19. On peut par exemple citer les outils fournis par le Ministère de l'Education Nationale et de la Jeunesse comme les Banques de Ressources Numériques Educatives (BRNE), ou bien les plateformes d'environnement virtuel comme Studio Code utilisé dans les classes pour apprendre les notions de la programmation, ou encore GraphoGame pour l'apprentissage de la lecture pour les enfants dyslexiques (RUIZ et al., 2017). Ce domaine en pleine expansion, également désigné sous le nom d'*EdTech* (WELLER, 2018), regroupe des entreprises spécialisées dans les nouvelles technologies et en éducation. En 2015, 80000 applications sur l'App Store d'Apple étaient des applications éducatives pour l'apprentissage, et 72% des applications les plus achetées étaient des applications éducatives pour jeunes enfants (0-8 ans). Cependant, ces applications sont largement non régulées et non testées (HIRSH-PASEK et al., 2015). Ces environnements virtuels peuvent contenir des personnages virtuels, qui ont un impact positif ou négatif sur l'apprentissage. Dans la recherche en informatique, un sous-domaine de l'intelligence artificielle consiste à créer et évaluer ces personnages virtuels dans des environnements virtuels d'apprentissage.

En effet, les environnements virtuels intègrent souvent des *agents virtuels pédagogiques*. Les agents virtuels pédagogiques sont une sorte particulière d'agent virtuel incarné. Un agent virtuel incarné, aussi appelé Agent Conversationnel Animé (ACA), représente une interface dans laquelle le système est représenté par une entité autonome (personnage humanoïde ou non, réaliste ou peu réaliste), où les informations peuvent être transmises aux utilisateurs par de multiples modalités (voix, gestes, expressions faciales) (CASSELL, 2001 ; LUGRIN et al., 2022). Les agents virtuels n'utilisent pas nécessairement une communication multimodale, et peuvent par exemple communiquer uniquement via du texte (E.-J. LEE, 2003). Plus particulièrement, les agents virtuels *pédagogiques* sont des ACA dans des environnements numériques utilisés pour améliorer l'apprentissage en milieu éducatif (MOHTADI et al., 2014). Leur but est d'améliorer l'expérience des utilisateurs à des fins pédagogiques, notamment en les

aidant et en les guidant pour améliorer leur performance et leur apprentissage dans une tâche (SOLIMAN & GUETL, 2013). Les agents virtuels pédagogiques peuvent avoir un style cartoon ou plus réaliste, humanoïde ou non. Dans cette thèse, nous visons plus particulièrement à créer des agents virtuels pédagogiques représentés par des personnages d'aspect humain, sans aller dans l'ultra-réalisme.

Les recherches sur les agents virtuels pédagogiques visent à les créer et à évaluer la perception des individus selon plusieurs caractéristiques, ainsi que l'efficacité de ces entités artificielles sur l'apprentissage dans un domaine particulier. C'est ainsi que des études ont montré que les agents pédagogiques peuvent augmenter la motivation des étudiants, réduire l'anxiété, et augmenter les interactions avec le système (ATKINSON, 2002 ; GULZ, 2004 ; R. MORENO, 2005). Bien que certaines études ont montré que les agents virtuels pédagogiques peuvent avoir des effets négatifs, notamment en augmentant la charge cognitive des utilisateurs (CHOI & CLARK, 2006 ; CLARK & CHOI, 2007), la plupart des études montrent que ces agents peuvent améliorer les performances des étudiants. D'après la méta-analyse de 43 études de SCHROEDER et al. (2013), les agents virtuels pédagogiques ont un effet positif sur l'apprentissage, notamment sur la rétention de connaissance, le rappel de connaissance, et le transfert de connaissance pour résoudre des problèmes non vus auparavant.

La simple présence d'un agent virtuel pédagogique peut ne pas suffire à améliorer l'apprentissage des utilisateurs (LI et al., 2016). Ainsi, dans l'étude de R. MORENO et al. (2001) sur une tâche de rappel des connaissances en biologie, les étudiants avec un agent virtuel pédagogique n'ont pas mieux performé que ceux sans agent. Cependant, dans l'étude de GRIVOKOSTOPOULOU et al. (2020), les étudiants ayant interagi avec un agent virtuel pédagogique en tant que compagnon d'apprentissage ont vu leurs résultats s'améliorer dans un test d'ingénierie environnementale. Effectivement, plusieurs aspects des agents virtuels pédagogiques sont à prendre en compte. L'effet sur l'apprentissage diffère suivant le rôle (Y. KIM & BAYLOR, 2016) et le niveau de réalisme (BAYLOR & KIM, 2004) des agents virtuels pédagogiques (Section 3.3). De plus, l'interaction socio-émotionnelle positive entre un utilisateur et un agent virtuel est étudié dans divers domaines pour améliorer l'expérience de l'utilisateur. Ce genre d'interaction positive, connu sous le nom de *rapport*, est atteignable par trois composantes : la positivité, l'attention mutuelle, et la coordination (TICKLE-DEGNEN & ROSENTHAL, 1990). Dans le domaine pédagogique, KRÄMER et al. (2016) ont ainsi montré que le rapport dans les interactions humain-agent a permis aux apprenants d'obtenir de meilleurs résultats en mathématiques.

Les agents virtuels pédagogiques peuvent donc permettre d'améliorer les performances des utilisateurs dans différentes tâches d'apprentissage. Cependant, certaines caractéristiques de ces agents, comme le genre, peuvent avoir une influence importante sur les effets sur l'apprentissage. On peut constater que peu d'attention a été accordée à cette caractéristique, alors que le genre a une influence importante sur l'apprentissage (Chapitre 2). Or, les études en sciences humaines et sociales sont

particulièrement inspirantes pour la création des agents virtuels, car les recherches montrent que les individus ont tendance à interagir avec les agents virtuels comme s'ils étaient humains (HEYSELAAR et al., 2023; NASS & MOON, 2000). On peut faire l'hypothèse que les environnements virtuels d'apprentissage n'échappent donc pas à la menace du stéréotype (Chapitre 2). ***Il est donc primordial d'explorer l'effet du genre des agents virtuels pour éviter que cela engendre des contre-performances, notamment si la menace du stéréotype n'est pas réduite dans ces environnements virtuels.***

En partant des résultats de recherche en cognition sociale (Chapitre 2) et dans le domaine des agents virtuels pédagogiques, on peut émettre l'hypothèse qu'un agent virtuel pédagogique permettrait de réduire les effets de menace du stéréotype, et pourrait ainsi être utilisé comme une aide supplémentaire en soutien à la pédagogie des enseignants et des parents. Cependant, à notre connaissance, peu d'étude a testé l'efficacité d'un agent virtuel pédagogique contre la menace du stéréotype. Cela pourrait s'expliquer par le peu de travaux entre le domaine sur la menace du stéréotype et celui des agents virtuels. Si l'on applique les résultats en cognition sociale de la menace du stéréotype concernant les filles et les femmes en mathématiques (Chapitre 2) pour la conception d'agents virtuels pédagogiques, il est important de : (1) créer l'apparence des agents virtuels selon les utilisateurs (âge, genre, origine culturelle, vêtements), (2) créer un discours rapprochant les agents virtuels d'un rôle de modèle de réussite pour que l'élève puisse s'identifier à eux, (3) donner la capacité aux agents virtuels d'interagir avec des comportements de rapport. La conception d'un agent virtuel pédagogique pouvant réduire la menace du stéréotype soulève plusieurs questions de recherche que l'on explore à travers un état de l'art dans ce chapitre, organisé comme suit :

- *Est-ce que le genre d'un agent virtuel a un impact sur la perception et la performance des utilisateurs ?*
  En effet, nous avons montré dans le chapitre 2 que le genre est important dans le processus d'apprentissage. Notamment, les filles s'identifient aux filles, et les garçons s'identifient aux garçons pour mieux apprendre. De plus, les élèves reportent moins de stéréotype de genre portant sur les compétences en mathématiques des femmes lorsque leur enseignant en mathématiques est une femme. Cette première section nous permettra d'explorer si le genre des agents virtuels a le même type d'impact sur les individus - **Section 3.2**.
- *Existe-t-il des agents incarnant un modèle de réussite social en mathématiques ou en sciences ?*
  Si de tels agents existent, on pourrait s'inspirer de ce type d'agent pour contrer la menace du stéréotype - **Section 3.3**.
- *Comment modéliser les interactions positives (e.g., le rapport), et cela a-t-il un impact sur les performances des utilisateurs ?*
  L'apprentissage étant un processus cognitif mais également social (BANDURA, 1986), le rapport permet d'améliorer l'apprentissage des élèves dans les études interhumaines (FRISBY & MARTIN, 2010; GRANITZ et al., 2009; E. MURPHY &

RODRÍGUEZ-MANZANARES, 2012). Le rapport *virtuel* lors de l'interaction entre un agent virtuel et un utilisateur pourrait avoir les mêmes effets. De plus, la composante de la coordination du rapport permettant une plus grande similitude entre les interactants, cela pourrait renforcer le niveau d'identification pour percevoir le tuteur comme un modèle de réussite - **Section 3.4**.

3.2. Impact du genre des agents virtuels sur la perception et la performance des utilisateurs

Nous présentons dans un premier temps une revue systématique des recherches sur les études perceptives des agents virtuels en fonction de leur genre, quel que soit le domaine d'application, et sur l'impact du genre des agents virtuels dans le cadre d'une tâche d'apprentissage. Sur la base d'une revue détaillée de 59 articles, cette revue systématique porte sur la façon dont les agents virtuels sont conçus, et l'impact de leur genre sur différents résultats scolaires (motivation, apprentissage, intérêt), mais aussi sur la perception des agents virtuels par les participants. Les questions de recherche guidant cette revue sont les suivantes :

1. Les individus perçoivent-ils différemment les agents virtuels selon leur propre genre et celui de l'agent ?
2. Le genre des agents virtuels pédagogiques influence-t-il les performances ainsi que d'autres facteurs éducatifs (motivation, sentiment de réussite, ...) des élèves ?
3. Y a-t-il des tâches ou des situations académiques auxquelles un agent virtuel masculin est mieux adapté qu'un agent virtuel féminin, et vice versa, selon des études empiriques ?
4. Comment le rôle d'un agent virtuel impacte-il ces résultats ?
5. Comment l'apparence et les capacités interactives d'un agent virtuel impactent-elles ces résultats ?
6. Les agents virtuels androgynes sont-ils une solution potentielle pour lutter contre les stéréotypes de genre ?

Nous avons utilisé la méthode PRISMA pour répondre à ces questions, que nous présentons en premier. Nous détaillons ensuite les évaluations utilisées dans les articles de cette revue systématique, pour mieux comprendre les prochaines sections sur la persistance des stéréotypes de genre dans les interactions humain-agent, et l'effet du genre des agents virtuels pédagogiques sur les résultats scolaires. Les conclusions tirées de cette revue sont ensuite discutées.

3.2.1. Méthode de la revue systématique

3.2.1.1. Stratégie de recherche d'articles

Nous examinons les recherches sur l'impact du genre des agents virtuels sur le comportement et les perceptions des utilisateurs, puis sur les apprenants. Pour cela,

nous avons examiné des articles de la base de données *Web of Science* sur une période de 21 ans allant de 2000 à 2021. Pour collecter les études pertinentes, nous avons effectué une recherche dans la base de données en ligne avec la requête *gender+("virtual agent*" OR "virtual character*")*. Cette revue systématique a été menée selon les directives PRISMA (WEBSTER & WATSON, 2002) comme suit : (1) en parcourant les bases de données et en commençant par les contributions majeures dans les principales revues, (2) en passant en revue les citations des articles identifiés à l'étape 1 pour déterminer les articles antérieurs qui devraient être pris en compte, et (3) en identifiant les articles qui citent les articles trouvés dans les étapes précédentes. Nous avons utilisé Google Scholar pour la dernière étape. Au total, 120 articles ont été retenus après avoir suivi ces étapes.

3.2.1.2. Sélection des articles

À partir de cet ensemble d'articles, nous avons sélectionné les études empiriques analysant l'effet du genre des agents virtuels sur les perceptions, les comportements et les résultats scolaires des utilisateurs. Nous n'avons pris en compte que les agents virtuels incarnés, c'est-à-dire que nous avons exclu les études sur les assistants vocaux. Nous nous sommes concentrés sur la culture occidentale et n'avons donc sélectionné que des articles relatifs à cette culture. Nous avons exclu les articles concernant les avatars (utilisateurs contrôlant un agent virtuel) qui concernaient principalement les jeux vidéo. Finalement, nous avons retenu un ensemble de 59 articles. Nous avons distingué deux types d'articles de recherche : les études perceptives d'agents virtuels selon leur genre, quel que soit le domaine d'application, et les études de recherche sur l'impact du genre des agents pédagogiques dans le cadre d'une tâche d'apprentissage.

3.2.2. Évaluations des agents virtuels

Nous présentons dans cette sous-section les évaluations utilisées dans les articles sélectionnés pour cette revue. Nous avons classé ces évaluations dans 3 groupes que nous présentons ci-dessous.

3.2.2.1. Évaluations de la perception des utilisateurs

La plupart des études ont utilisé des questionnaires post-expérience pour évaluer les perceptions des utilisateurs envers les agents virtuels. Le plus connu et utilisé est sans doute celui du Godspeed de BARTNECK et al. (2009). Ce questionnaire comprend les échelles d'intelligence, d'anthropomorphisme, de sympathie, d'animation, et de sentiment de sécurité avec l'agent virtuel. Récemment, FITRIANIE et al. (2020) ont constaté que les chercheurs dans le domaine des agents virtuels créaient des questionnaires pour évaluer leurs agents virtuels, à la place d'utiliser des questionnaires existants et validés. À partir de 81 études, ils ont identifié 189 échelles utilisées dans 89 questionnaires, mesurant parfois les mêmes aspects. Les auteurs ont unifié ces échelles en 19 ensembles (e.g., la confiance en l'agent, la personnalité de l'agent)

pouvant être utilisées par la communauté.

Concernant les perceptions liées au genre des agents, des éléments évalués sur une échelle de Likert ont été utilisés pour déterminer les attributions stéréotypées des agents par les participants. Elles correspondaient à des traits stéréotypés associés aux femmes (e.g., affectueux, compatissant, sensible et serviable), des traits stéréotypés associés aux hommes (e.g., arrogant, ambitieux, agressif, courageux et décisif), et des traits de compétence plus souvent associés aux hommes (e.g., cultivé, intelligent, expert, crédible, créatif, innovant et organisé) (FENG et al., 2017; KANTHARAJU et al., 2018; KHASHE et al., 2017; E.-J. LEE, 2003; NUNAMAKER et al., 2011; SCZESNY et al., 2018b; van der LUBBE & BOSSE, 2017). Un autre questionnaire est parfois utilisé pour déterminer quels traits stéréotypés les utilisateurs appliquent aux agents (KULMS et al., 2011). Cette échelle, appelée le *Bem Sex Role Inventory* (BSRI) et développée par BEM (1974) en sciences sociales, mesure la construction du schéma de genre des individus, vise à mettre en évidence l'androgynie, et interroge la dichotomie habituelle des traits genrés féminins/masculins attribués stéréotypiquement aux personnes. Le BSRI se compose de 20 éléments positifs stéréotypiquement associés aux hommes (e.g., indépendant, analytique), 20 autres éléments positifs stéréotypiquement associés aux femmes (e.g., compatissant, aime les enfants) et 20 autres éléments positifs neutres (e.g., plein de tact, fiable). Enfin, dans la littérature, la perception du genre des agents est parfois évaluée simplement à l'aide d'une échelle de Likert à 7 points, avec homme=1, androgyne=4, et femme=7 (E.-J. LEE, 2003; MCDONNELL et al., 2009; NAG & YALÇIN, 2020; NICULESCU et al., 2009). Il est intéressant de noter que ces questionnaires n'ont pas été créés pour évaluer spécifiquement les agents virtuels. Les questionnaires des études interhumaines peuvent être adaptés voire utilisés pour évaluer des agents virtuels.

D'autres perceptions d'attitude sociale ont également été évaluées avec des éléments sur une échelle de Likert, comme la convivialité perçue, la fiabilité, ou encore la présence sociale de l'agent (AKBAR et al., 2018; GUADAGNO et al., 2007; KHASHE et al., 2017; E.-J. LEE, 2003; LUNARDO, BRESSOLLES et al., 2016; NUNAMAKER et al., 2011). La présence sociale est communément définie comme la sensation d'être en présence d'une personne réelle (BIOCCA, 1997). Elle peut être évaluée avec le *Networked Minds Questionnaire* (e.g., « L'autre individu ne m'a pas remarqué dans la pièce ») (BIOCCA et al., 2001), tel qu'utilisé par KULMS et al. (2011).

Dans cette thèse, nous avons utilisé le questionnaire du Godspeed (BARTNECK et al., 2009) qui correspond aux échelles numéro 1 et 4 de FITRIANIE et al. (2020) pour évaluer respectivement l'anthropomorphisme et la sympathie de nos agents virtuels pédagogiques (Chapitre 7). Cela nous a permis d'explorer les éventuels stéréotypes de genre que les élèves peuvent avoir envers nos agents virtuels pédagogiques (e.g., différence de niveau d'anthropomorphisme perçu selon le genre et les gestes des agents virtuels, différence de sympathie perçue selon le genre des agents virtuels). Le niveau d'anthropomorphisme perçu nous a également permis de se situer au niveau

de la *vallée de l'étrange*, phénomène poussant les utilisateurs à rejeter l'agent virtuel lorsque la précision de ses mouvements est en décalage avec son niveau de réalisme physique (KÄTSYRI et al., 2015).

3.2.2.2. Évaluations objectives de l'apprentissage des utilisateurs

L'impact des agents pédagogiques sur l'apprentissage des utilisateurs peut être évalué en mesurant la performance des utilisateurs dans un exercice en comparant différentes conditions : par exemple, des agents virtuels avec des comportements différents (CHANG et al., 2019), ou encore la présence d'agents virtuels genrés (Y. KIM & WEI, 2011). Les performances peuvent être analysées à l'aide de tests de compétence pour, par exemple, mesurer la rétention des connaissances (utiliser des connaissances acquises pour résoudre un problème, SAJJADI et al., 2020), le rappel (la capacité à se souvenir des éléments, WIRZBERGER et al., 2019) ou l'apprentissage par transfert (en utilisant des connaissances acquises pour résoudre de nouveaux problèmes, MAKRANSKY et al., 2019). En plus des performances, les chercheurs peuvent également évaluer le temps de réponse nécessaire pour résoudre un problème (HAYES et al., 2010), ainsi que l'effort des élèves, se mesurant par le nombre de problèmes résolus (qui ne sont pas nécessairement corrects) lors d'un test de compétence (KRÄMER et al., 2016).

Dans cette thèse, nous avons mesuré la performance des élèves via des mesures objectives comme la rétention de connaissances lors d'un test de mathématiques standardisé, après avoir interagi avec un de nos agents virtuels pédagogiques (Chapitre 9).

3.2.2.3. Évaluations subjectives de l'apprentissage des utilisateurs

Dans les situations d'apprentissage, d'autres mesures plus subjectives que la performance peuvent être notées à l'aide d'éléments sur une échelle de Likert. Ces mesures comprennent l'intérêt pour une tâche ou un domaine (e.g., « Je vais choisir de suivre un cours de sciences dures. », R. ROSENBERG-KIMA et al., 2008), les croyances sur l'utilité d'une tâche ou d'un domaine (e.g., « J'aurais beaucoup de bonnes opportunités de carrière si j'étais étudiant en sciences dures. », PLANT et al., 2009), l'auto-efficacité des apprenants comme le fait de se sentir capable d'accomplir une tâche (e.g., « Je peux obtenir de bonnes notes en mathématiques. », Y. KIM et WEI, 2011), l'autorégulation des apprenants notamment pour réguler leurs comportements afin de réussir une tâche (e.g., « J'ai suivi ma progression. », BAYLOR et KIM, 2004), la motivation des apprenants évaluée avec l'échelle du *Situational Motivation Scale* (SIMS) (GUAY et al., 2000) qui comprend 16 items sur la motivation à travailler sur des tâches (e.g., « Parce que je le fais pour mon propre bien. », KRÄMER et al., 2016), le plaisir des apprenants (e.g., « À quel point avez-vous apprécié réviser pour l'examen ? », SHIBAN et al., 2015), leur efficacité d'apprentissage perçue (e.g., « J'ai acquis une bonne compréhension des concepts de base. », SAJJADI et al., 2020), et leur exigence mentale pour savoir

combien de ressource mentale leur a été nécessaire pour par exemple penser, décider, ou calculer (e.g., « La tâche était-elle facile ou exigeante ? », HART et STAVELAND, 1988 ; PEZZULLO et al., 2017).

Dans cette thèse, avec ces échelles de mesure, nous avons évalué l'importance des mathématiques et le sentiment de réussite (e.g., l'auto-efficacité) en mathématiques des élèves, avant et après avoir interagi avec un de nos agents virtuels pédagogiques (Chapitre 9).

3.2.3. Persistance des stéréotypes de genre dans les interactions humain-agent

Dans cette section, nous nous intéressons à la question des stéréotypes associés aux agents virtuels, quel que soit le domaine. Nous cherchons à savoir si les agents virtuels sont perçus différemment selon leur apparence genrée. Nous avons résumé les études sélectionnées sur les perceptions des utilisateurs des agents virtuels en fonction de leur genre dans chaque ligne du Tableau 3.1. Nous avons indiqué les caractéristiques des agents, le nombre de participants hommes et femmes avec leur âge moyen, les tâches de l'étude, les mesures observées et les résultats de l'étude.

TABLEAU 3.1

Synthèse d'articles sur les études perceptives d'agents virtuels en fonction de leur genre, quel que soit le domaine d'application. Les articles sont classés du plus ancien au plus récent. AF = Agent(s) Femme(s), AH = Agent(s) Homme(s), PF = Participante(s) Femme(s), PH = Participant(s) Homme(s). La colonne des agents décrit le nombre d'agents en fonction de leur genre, leur dimension (2D ou 3D), leur apparence (réaliste ou cartoon) et leur rôle. La colonne des participants décrit le nombre d'hommes et de femmes qui ont participé à l'étude, ainsi que l'âge moyen. Dans la colonne des résultats, par exemple, "PH > PF" signifie que cela a eu plus d'impact sur les participants hommes que sur les participantes. "AF > AH" signifie que l'agent femme a eu plus d'impact que l'agent homme sur la mesure en question. Les explications se trouvent après ce tableau.

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
E.-J. LEE (2003)	• 1 AH • 1 AF • 2D • Textee • Cartoon • Conseiller	• 28 PH • 88 PF • avg age N/A	Jouer à un jeu à choix multiples avec un agent. Les participants pouvaient modifier leur réponse après avoir reçu la réponse de l'agent. Il a été précisé que la réponse de l'agent pourrait ne pas être correcte.	• Masculinité • Attraction • Compétence • Fiabilité • Persuasion (questions sur le sport ou les cosmétiques)	• *Masculinité* : AH > AF • *Attraction : compétence* : AF > AH • *Persuasion (sport)* : AH > AF • *Persuasion (cosmétique)* : AF > AH
ZANBAKA et al. (2006)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Orateur	• 41 PH • 97 PF • avg age 20.6	Écouter les agents livrer un message pour changer les attitudes des participants à l'égard des examens à l'université.	• Persuasion	• *Persuasion* : - PH : AF > AH - PF : AH > AF

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
GUADAGNO et al. (2007)	• 1 AH • 1 AF • 1 neutre • 3D • Voix • Réaliste • Orateur	• 37 PH • 29 PF • avg age N/A	Écouter les agents parler des changements apportés à la politique de sécurité de l'université.	• Sympathie • Crédibilité • Qualité de la présentation • Persuasion	• *Sympathie, crédibilité, qualité de la présentation* : - PH : pas significatif - PF : AF > AH • *Persuasion* : - PH : AH > AF - PF : AF > AH
GUADAGNO et al. (2007)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Orateur	• 85 PH • 89 PF • avg age N/A	Écouter les agents parler des changements apportés à la politique de sécurité de l'université.	• Sympathie • Crédibilité • Qualité de la présentation • Présence social • Persuasion	• *Sympathie* : AF > AH • *Crédibilité, qualité de la présentation* : pas significatif • *Persuasion* : - PH : AH > AF - PF : pas significatif

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
GULZ et al. (2007)	• 2 AH • 2 AF • 2D • Voix • Réaliste • Présentateur	• 72 PH • 86 PF • avg age N/A	Ecouter des agents présentant les programmes d'ingénierie de l'université.	• Agent préféré • Intérêt	• *Agent préféré* : - PH : agents moins féminin & moins masculin > agent plus masculin > agent plus féminin - PF : agents moins féminin & moins masculin > agent plus féminin > agent plus masculin • *Intérêt* : - PH : agents plus féminin & plus masculin > agents moins féminin & moins masculin - PF : agents plus féminin, & plus masculin, & moins féminin > agent moins masculin
DILL et al. (2008)	• 16 AH • 16 AF • 3D • Réaliste • Personnage de jeux vidéo	• 61 PH • 120 PF • avg age 18.82	Regarder une présentation PowerPoint opposant des images fixes de personnages de jeux vidéo et des photos de sénateurs américains (hommes ou femmes). Lecture d'une histoire vraie sur le harcèlement sexuel d'une étudiante par un professeur.	• Tolérance pour le harcèlement sexuel • Attitude favorable au viol	• *Tolérance pour le harcèlement sexuel* : PH > PF • *Attitude favorable au viol* : PH > PF

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
R. ROSENBERG-KIMA et al. (2008)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Orateur	• 89 PF • avg age 19.7	Écouter l'histoire de 4 femmes ingénieures et les avantages de l'ingénierie, expliquer par un agent ou par une voix.	• Intérêt • Auto-efficacité • Utilité pour l'ingénierie	• *Intérêt, auto-efficacité, utilité pour l'ingénierie* : pas significatif
R. ROSENBERG-KIMA et al. (2008)	• 4 AH • 4 AF • 3D • Voix • Réaliste • Orateur	• 111 PF • avg age 19.72	Écouter l'histoire de 4 femmes ingénieures et les avantages de l'ingénierie, expliquer par un agent ou par une voix.	• Utilité pour l'ingénierie • Moins de stéréotypes de genre en ingénierie	• *Utilité pour l'ingénierie* : AH > AF (pas significatif) • *Moins de stéréotypes de genre en ingénierie* : AF > AH
NICULESCU et al. (2009)	• 3 AH • 3 AF • 1 neutre • 3D • Voix • Cartoon • Assistant	• 24 PH • 24 PF • avg age N/A	Interagir avec des agents à propos de questions médicales, évaluer le genre d'un agent androgyne après ou avant d'évaluer les agents non androgynes.	• Genre perçu de l'agent androgyne	• *Genre de l'agent androgyne* : - Agents non-androgynes évalués en 1er : - PF : plus féminine - PH : plus masculin - Agent androgyne évalué en 1er : - PF : plus masculin - PH : plus féminin

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
McDonnell et al. (2009)	• 1 AH • 1 AF • 2 neutre • 3D • Réaliste • Sujet	• 22 PH • 19 PF • avg age N/A	Regarder une vidéo d'agents qui marchent.	• Genre perçu des agents	• *Genre perçu des agents* : - AF (marche masculine) : ambiguë - AF (marche neutre) : femme - AH (marche féminine) : ambiguë - AH (marche neutre) : homme - agents neutres : ambiguë - agents neutres (marche féminine) : femme - agents neutres (marche masculine) : homme - agents neutres (marche neutre) : femme
McDonnell et al. (2009)	• 3 AH • 3 AF • 3D • Réaliste • Sujet	• 33 PH • 5 PF • avg age N/A	Regarder une vidéo d'agents qui marchent.	• Genre perçu des agents	• *Genre perçu des agents* : - AF évalué 'plus femme' : AF (hanches et taille de poitrine plus grandes) > AF (hanches et taille de poitrine plus petites) - AH évalué 'plus homme' : pas de différence - Agents évalués 'plus ambiguë' : AF (marche masculine) & AH (marche féminine)

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
FOX et BAILENSON (2009)	• 4 AF • 3D • Réaliste • Sujet	• 43 PH • 40 PF • avg age 20.82	Rencontrer un agent (évite le regard (Non-Regard) ou soutien le regard (Regard), vêtements masculins ou féminins) via la réalité virtuelle, émettre des jugements à son sujet.	• Acceptation du mythe du viol • Sexisme bienveillant • Sexisme hostile	• *Acceptation du mythe du viol* : agent masculine Non-Regard > agent féminine Regard > agent masculine Regard > agent féminine Non-Regard • *Sexisme bienveillant* : agent masculine Non-Regard > agent féminine Non-Regard > agent masculine Regard • *Sexisme bienveillant* : agents Non-Regard > agents Regard • *Sexisme hostile* : agent féminine Regard > agent masculine Regard
CLOUD-BUCKNER et al. (2009)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Guide	• 19 PH • 16 PF • avg age N/A	Regarder un agent présenter un campus universitaire en tant que guide touristique en ligne.	• Convivialité • Colère • Coopération • Conscience de soi • Sens de l'aventure • Sympathie • Sociabilité • Assurance • Coopération • Autodiscipline	• *Convivialité, colère, coopération, conscience de soi, sens de l'aventure, sympathie* : Personnalité extravertie : AH > AF • *sociabilité, assurance, coopération, conscience de soi, autodiscipline* : Personnalité introvertie : AF > AH

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
NICULESCU et al. (2010)	• 1 AH • 1 AF • 1 neutre • 3D • Textee • Cartoon • Tuteur	• 4 PH • 4 PF • avg age N/A	Poser des questions médicales à un agent.	• A l'aise • Confiant • Moins tendu • Agent préféré	• *A l'aise, confiant, moins tendu* : AF > AH &agent androgyne • *Agent préféré* : AF > AH > agent androgyne
R. B. ROSENBERG-KIMA et al. (2010)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Orateur	• 119 PF • avg age 21.49	Écouter un agent décrire quatre femmes ingénieures et les avantages de l'ingénierie.	• Intérêt • Auto-efficacité • Utilité • Sympathie • Moins de stéréotypes de genre en ingénierie	• *Intérêt* : - PF Noires : AF Noire > autres - PF Blanches : AF > AH • *Auto-efficacité, utilité, sympathie* : - PF Noires (agents Noirs) : AF > AH - PF Noires (agents Blancs) : AH > AF • *Moins de stéréotypes de genre en ingénierie* : - PF Noires : agents Noirs > agents Blancs - PF Blanches : AF > AH
ASTRID et al. (2010)	• 1 AF • 3D • Voix • Réaliste • Interrogateur	• 41 PH • 42 PF • avg age 37.27	Répondre aux questions personnelles d'un agent.	• Faible • Timide • Naïf • Compatissant • Attrayant	• *Faible, timide, naïf, compatissant, attrayant* : pas significatif

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
NUNAMAKER et al. (2011)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Interrogateur	• 53 PH • 35 PF • avg age 25.45	Répondre aux questions d'un agent simulant un contrôle à l'aéroport.	• Pouvoir • Fiabilité • Expertise • Sympathie	• *Pouvoir, fiabilité, expertise* : AH > AF • *Sympathie* : AF > AH
KULMS et al. (2011)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Interrogateur	• 32 PH • 40 PF • avg age 35.03	Répondre à des questions informelles posées par un agent, condition évite le regard (Non-Regard) ou soutien le regard (Regard).	• Masculinité • Évaluation positive • Présence sociale	• *Masculinité* : AH Regard > AH Non-Regard • *Évaluation positive* : AF > AH • *Présence sociale* : AH > AF
BRAHNAM et DE ANGELI (2012)	• 8 AH • 8 AF • 3 neutre • 2D • Textee • Cartoon • Chatbot	• 127 PH • 73 PF • avg age N/A	Discuter par message avec un chatbot.	• Discours sexuel • Nombre moyen de mots à propos d'argents/de travail, et d'apparence physique	• *Discours sexuel, Nombre moyen de mots (apparence physique)* : AF > AH • *Nombre moyen de mots (argents/travail)* : AH > AF (agents adultes)
OZOGUL et al. (2013)	• 2 AH • 2 AF • 3D • Cartoon • Tuteur	• 35 PH • 42 PF • avg age 12.83	Noter les photos des agents.	• Agent préféré • Agent préféré pour apprendre l'ingénierie	• *Agent préféré* : - PF : AF > AH - PH : AH > AF • *Agent préféré pour apprendre l'ingénierie* : jeune AF > jeune AH > vieil AH > vieil AF

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
PAYNE et al. (2013)	• 4 AH • 4 AF • 2D & 3D • Cartoon • Assistant	• 220 PH • 358 PF • avg age 35.56	Choisir un agent pour aider aux caisses libre-service.	• Agent préféré	• *Agent préféré* : - PF : AF > AH - PH : AH > AF
LUNARDO, BRESSOLLES et al. (2016)	• 2 AH • 2 AF • 2D • Textee • Réaliste • Assistant	• 107 PH • 147 PF • avg age N/A	Interagir avec un agent par message sur fnac.com.	• Attraction	• *Attraction* : - Agents (vêtements professionnels) : AF > AH - Agents (vêtements décontractés) : AF > AH (pas significatif)
van der LUBBE et BOSSE (2017)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Employé	• 55 PH • 38 PF • avg age N/A	Interagir avec un agent employé pour négocier le salaire de l'agent (agent assertif ou agent non-assertif).	• Langage approprié • Sensibilité • Aucun accord conclu • Persuasion	• *Langage approprié* : AF assertive > AH assertif • *Sensibilité* : AH non-assertif > AF non-assertive • *Aucun accord conclu* : AH assertif > AF assertive > AF non-assertive > AH non-assertif (pas significatif) • *Persuasion* : AF assertive > AH assertif
FENG et al. (2017)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Instructeur	• 31 PH • 32 PF • avg age 21.37	Rejouer une scène en présence d'un agent donnant un feedback négatif.	• Inspiration • S'auto-blâmer • Aide • Agent préféré	• *Inspiration, s'auto-blâmer, aide, agent préféré* : AF > AH

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
MELL et al. (2017)	• 1 AF • 2D • Textee • Réaliste • Assistant	• 241 PH • 140 PF • avg age 35.13	Répondre aux questions d'un chatbot sur des informations sensibles, soit avec une photo d'une vraie femme (humain), soit avec une photo d'un agent virtuel féminin (AF), soit sans photo (non-présence).	• Mensonges signalés • Autoriser le système à effectuer une vérification de crédit • Donner son adresse	• *Mensonges signalés* : humain > non-présence > AF • *Autoriser le système à effectuer une vérification de crédit* : - PF : AF > humain > non-présence - PH : non-présence > AF > humain • *Donner son adresse* : - PF : humain = AF = non-présence - PH : humain > AF > non-présence
KHASHE et al. (2017)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Orateur	• 98 PH • 116 PF • avg age N/A	Être demandé pour éteindre les lumières et ouvrir la fenêtre par un responsable. Le/la responsable peut être soit une voix uniquement (masculine ou féminine), soit un texte uniquement (d'une femme ou d'un homme), ou un agent virtuel (femme ou homme).	• Affectueux • Amical • Sympathie • Persuasion	• *Affectueux, amical, sympathie* : AF et voix féminine > AH et voix masculine • *Persuasion* : AF, voix féminine, texte d'une femme > AH, voix masculine, texte d'un homme

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
KANTHARAJU et al. (2018)	• 2 AH • 2 AF • 3D • Voix • Réaliste • 2 experts • 2 motivateurs	• 113 PH • 92 PF • avg age N/A	Écouter une conversation persuasive sur le cinéma entre agents.	• Distant • Arrogant • Énergique • Crédible • Persuasion	• *Distant, arrogant, énergique, crédible, Persuasion* : PH > PF
AKBAR et al. (2018)	• 1 AH • 1 AF • 3D • Textee • Réaliste • Interviewer	• 158 PH • 158 PF • avg age N/A	Être interviewé par un agent par message pour un emploi dans une société financière.	• Amabilité • Fiabilité	• *Amabilité* : agent du genre opposé aux participants > agent du même genre que les participants • *Fiabilité* : agent du même genre que les participants > agent du genre opposé aux participants
MOUSAS et al. (2018)	• 2 AH • 3D • Réaliste • Sujet	• 56 PH • 16 PF • avg age 23.24	Répondre aux questions sur les agents (e.g., « Seriez-vous mal à l'aise si ce personnage virtuel communiquait avec vous ? ») par l'expérimentateur pendant que l'agent se dirigeait vers le participant.	• A l'aise • Prêt à interagir • Sympathie	• *A l'aise, Prêt à interagir* : PH > PF • *Sympathie* : - agent zombie : PH > PF - AH : pas significatif

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
AIT CHALLAL et GRYNSZPAN (2018)	• 1 AH • 1 AF • 3D • Réaliste • Sujet	• 12 PH • 12 PF • avg age 23.6	Regarder des agents virtuels assis devant (dans les conditions de : suivi du regard, évitement du regard, regard direct élevé et de regard direct faible). Juger leurs personnalités.	• Névrosisme • Amabilité	• *Névrosisme* : AF > AH • *Amabilité (regard direct élevé)* : AH > AF
ter STAL et al. (2020)	• 4 AH • 4 AF • 2D • Cartoon • 4 experts, 4 pairs	• 67 PH • 69 PF • avg age 51.36	Observer et noter 8 agents.	• Convivialité • Expertise • Autorité	• *Convivialité* : AF > AH • *Expertise, autorité* : AH > AF
ter STAL et al. (2020)	• 4 AH • 4 AF • 2D • Cartoon • 4 experts, 4 pairs	• 35 PH • 30 PF • avg age 67.85	Observer et noter 8 agents.	• Convivialité • Autorité	• *Convivialité* : pas significatif • *Autorité* : AH > AF
ZIBREK et al. (2020)	• 2 neutre • 3D • Réaliste • Sujet	• 10 PH • 10 PF • avg age N/A	Appuyer sur un bouton dès qu'on se sentait mal à l'aise avec la distance entre soi-même et un agent marchant vers nous.	• Genre perçu de l'agent neutre	• *Genre perçu de l'agent neutre* : - mouvements féminins = femme - mouvements masculins = homme

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
D. RICHARDS et al. (2020)	• 6 AH • 6 AF • 3D • Voix • Réaliste • Assistant	• 43 PH • 146 PF • avg age 21.7	Regarder 12 vidéos de 12 agents différents se présentant.	• Agent préféré (avant et après avoir regardé les vidéos)	• *Agent préféré (avant)* : - PF : le genre n'a pas d'importance > AF > AH - PH : le genre n'a pas d'importance > AH > AF • *Agent préféré (après)* : AF méditerranéenne > AF asiatique > AF Blanche
NAG et YALÇIN (2020)	• 1 AH • 1 AF • 1 neutre • 3D • Réaliste • Sujet	• 41 PH • 31 PF • avg age 21.7	Regarder des photos d'agents et les évaluer.	• Communalité • Agentivité • Compétence	• *Communalité* : AF > AH (pas significatif) • *Agentivité, compétence* : pas significatif
ESPOSITO et al. (2021)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Assistant	• 22 PH • 24 PF • avg age 71.59	Regarder une vidéo d'un agent parlant d'EHPAD.	• Volonté d'interagir avec l'agent • Attraction • Utilité • Présentable • Professionnel • De bon goût • Agréable • Original • Créatif • Captivant	• *Volonté d'interagir avec l'agent, attraction, utilité, présentable, professionnel, de bon goût, agréable, original, créatif, captivant* : AF > AH

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
ESPOSITO et al. (2021)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Assistant	• 20 PH • 25 PF • avg age 71.22	Regarder une vidéo d'un agent parlant d'EHPAD. (étude 2).	• Volonté d'interagir avec l'agent • Attraction • Utilité • Présentable • Professionnel • De bon goût • Agréable • Original • Créatif • Captivant	• *Présentable, professionnel, de bon goût, agréable* : AF > AH • *Volonté d'interagir avec l'agent, attraction, utilité, original, créatif, captivant* : pas significatif
VILARO et al. (2021)	• 3 AF • 3D • Voix, text • Réaliste • Assistant, expert	• 53 PF • avg age 60.90	Regarder un agent délivrer des messages de dépistage du cancer colorectal.	• Fiabilité • Expertise	• *Fiabilité* : pas significatif • *Expertise* : agents (blouse médicale blanche) > agent (vêtements décontractés)
ANTONIO GÓMEZ-JÁUREGUI et al. (2021)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Interviewer	• 16 PH • 16 PF • avg age 29.95	Se présenter à un agent virtuel au visage flouté pour un entretien d'embauche.	• Dominance • Chaleureux	• *Dominance* : pas significatif • *Chaleureux* : AF (mouvements imités) > AF (mouvements aléatoires)

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
ŚWIDRAK et al. (2021)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Joueur	• 15 PH • 19 PF • avg age 25	Jouer à un jeu de négociation/prise de décision avec un agent féminin et un agent masculin.	• Touché agréable • Touché gênant • Touché adéquat • Persuasion	• *Touché agréable* : AF > AH • *Touché gênant* : PF > PH • *Touché adéquat* : AF plus masculine > AF moins masculine • *Persuasion* : - PH : agents plus masculins > agents moins masculins - PF : dépends de l'offre
ŚWIDRAK et al. (2021)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Joueur	• 40 PH • avg age 23	Jouer à un jeu de négociation/prise de décision avec un agent féminin et un agent masculin.	• Masculinité • Touché agréable • Touché gênant • Touché adéquat • Persuasion	• *Masculinité* : AF masculine > AH féminin • *Touché agréable* : AF > AH féminin • *Touché gênant* : AH féminin > autres (pas significatif) • *Touché adéquat* : autres > AH féminin (pas significatif) • *Persuasion* : agents perçus comme masculins > agents perçus comme féminins

Impact sur les sujets de conversations. Ces études montrent que les stéréotypes de genre persistent dans les interactions humain-agent. Le comportement des utilisateurs varie en fonction de l'apparence genrée des agents virtuels. Par exemple, dans une étude de DE ANGELI et BRAHNAM (2006), l'agent virtuel féminin a reçu un certain nombre de propositions sexuelles violentes allant jusqu'aux menaces de viol; l'agent virtuel masculin n'a reçu que quelques propositions sexuelles, aucune d'entre elles violentes ("**presse doucement mes lèvres contre les tiennes dans un petit baiser**"), et les autres commentaires sexuels émit lors des interactions avec l'agent virtuel masculin visaient son éventuelle petite amie. Dans une étude similaire menée par BRAHNAM et DE ANGELI (2012), les utilisateurs ont interagi avec plusieurs paires d'agents virtuels féminins/masculins, comprenant des agents enfants, des agents âgés, des agents Blancs, et des agents Noirs. Les agents virtuels féminins ont été la cible de beaucoup plus de discours sexuels, de commentaires sur leur apparence et d'insultes que les agents masculins; c'était également vrai pour la paire d'agents enfants. D'autres caractéristiques des agents virtuels ont influencé les sujets de conversation, telles que leur âge et leur apparence : les utilisateurs ont davantage parlé de travail, de réussite et d'argent avec des agents d'apparence plus âgée et vêtus de costumes qu'avec n'importe quelle autre paire d'agents. Cependant, les stéréotypes de genre s'appliquaient tout de même à cette catégorie, puisque les utilisateurs interagissaient davantage avec l'agent masculin plus âgé sur ces sujets qu'avec l'agent féminin plus âgé.

Impact sur les décisions des utilisateurs. Le genre d'un agent virtuel a également une influence directe sur les décisions des utilisateurs. Par exemple, E.-J. LEE (2003) a montré que les participants suivaient davantage les conseils des agents virtuels lorsque le genre des agents correspondait de manière stéréotypée au sujet (un agent féminin et les cosmétiques, un agent masculin et le sport). Ceci étant dit, dans cette étude, l'agent virtuel féminin avait une apparence particulièrement féminine, alors que l'agent masculin avait un aspect sportif. Ce résultat devrait donc être vérifié dans une étude séparée, par exemple en utilisant un agent virtuel féminin présentant une apparence sportive pour déterminer si ces résultats sont dus uniquement au genre de l'agent, et non à l'apparence globale des agents (vêtements et maquillage). Dans l'étude de GUADAGNO et al. (2007), l'agent virtuel masculin était plus persuasif lorsqu'il était perçu comme contrôlé par un ordinateur plutôt que contrôlé par un humain. L'inverse était vrai pour l'agent virtuel féminin. Les auteurs ont conclu que ces résultats peuvent avoir été dus à des stéréotypes de genre, en particulier par le fait que « les attentes des participants concernant l'interaction avec un ordinateur sont plus cohérentes avec les stéréotypes masculins (e.g., être compétent), alors que les attentes concernant l'interaction avec un humain sont plus cohérentes avec les stéréotypes féminins (e.g., être chaleureux) ». Non seulement le genre d'un agent virtuel peut influencer les décisions des participants, mais également son niveau de masculinité perçu. Dans un jeu de prise de décision où des agents virtuels faisaient une offre monétaire à des hommes participants, le nombre d'offres acceptées était plus élevé avec les agents perçus comme plus masculins (ŚWIDRAK et al., 2021). Les mêmes résultats ont été obtenus dans leur deuxième étude pour d'autres hommes

participants ; en revanche, les participantes ont accepté plus d'offres des agents que les hommes, mais n'ont été influencées que par l'offre en elle-même, et non par la masculinité perçue des agents.

Impact sur les traits associés aux agents. Le genre d'un agent a également un impact sur la façon dont les utilisateurs le perçoivent en termes de traits stéréotypés attribués aux hommes et aux femmes (SCZESNY et al., 2018b). Dans une étude de NUNAMAKER et al. (2011), l'agent masculin était perçu comme plus puissant, tandis que l'agent féminin était perçu comme plus sympathique. Même lorsque les agents masculins et féminins portaient les mêmes vêtements, présentaient les mêmes comportements verbaux et non-verbaux et avaient le visage flouté, l'agent féminin était noté comme plus chaleureux que l'agent masculin ; cependant, ils ont été évalués de la même manière pour la dominance, un trait généralement associé aux hommes (ANTONIO GÓMEZ-JÁUREGUI et al., 2021). En revanche, KULMS et al. (2011) ont constaté dans leur expérience principale que les participants n'attribuaient pas plus de traits masculins aux agents masculins ni plus de traits féminins aux agents féminins, contrairement à leur étude pilote avec 14 participants utilisant des images fixes des mêmes agents virtuels. Les auteurs ont conclu que les attributions stéréotypées devenaient moins importantes lorsque les participants pouvaient interpréter le comportement des agents. Cependant, une étude de AIT CHALLAL et GRYNSZPAN (2018) contredit cette conclusion : l'agent féminin était jugé moins agréable que l'agent masculin dans la condition où les agents regardaient directement l'utilisateur. Les auteurs ont suggéré que les participants étaient moins tolérants à la dominance lorsqu'elle était exprimée par un agent féminin. Les stéréotypes de genre associés au genre des utilisateurs peuvent également avoir un impact sur l'évaluation des agents virtuels, et donc sur les résultats des études. Dans une étude de MOUSAS et al. (2018), les hommes participants ont déclaré se sentir plus à l'aise avec un agent zombie que les participantes ; ils ont également plus aimé l'agent zombie que les participantes. Les auteurs ont conclu que les stéréotypes de genre pouvaient avoir influencé les résultats, car les stéréotypes appellent les hommes à être plus calmes face à la peur et à l'embarras/au dégoût, et à moins signaler des réactions émotionnelles liées à la peur. Il semblerait que les traits stéréotypés de genre sont attribués aux agents virtuels selon leur rôle et leur comportement, couplés à leur apparence genrée.

Effet du contexte stéréotypé. Les contextes associés de manière stéréotypée à un genre peuvent également avoir un impact sur les préférences des participants quant au genre des agents virtuels. Dans deux expériences menées par ter STAL et al. (2020), les participants âgés ont préféré des images fixes d'agents féminins dans un contexte médical. Selon les auteurs, ce résultat pourrait être dû au fait que le contexte est associé aux stéréotypes de genre féminin. De plus, les agents masculins ont été jugés plus autoritaires et experts que les agents féminins. Dans une étude de GULZ et al. (2007), lorsque des agents virtuels présentaient le programme universitaire en génie informatique, l'intérêt des participants était plus élevé avec les agents féminins et masculins par rapport aux agents d'apparence plus neutres (un agent femme moins

féminin, et un agent homme moins masculin). Cependant, les participants qui ont classé l'agent femme moins féminine comme étant la meilleure présentatrice l'ont choisi parce qu'ils pensaient qu'elle pouvait intéresser plus de filles à l'ingénierie informatique (« *elle a l'air jeune et gentille, et je pense qu'elle intéresserait plus de filles* ») ; et les participants qui ont classé l'agent femme féminine comme la pire présentatrice l'ont choisi parce qu'elle était une femme qui ne semblait pas appartenir à ce contexte (« *comme je l'ai dit, une femme est plus que bienvenue, mais elle avait l'air trop fashion, ce que je n'aime pas* »). Ces résultats montrent que les stéréotypes de genre s'appliquent à l'apparence des agents.

Effet du rôle stéréotypé. Outre le contexte, les rôles des agents virtuels peuvent également influencer la façon dont les utilisateurs les perçoivent. Lorsque les agents féminins ont été présentés en tant qu'assistants à des personnes âgées dans un contexte médical, les participants ont préféré interagir avec ces agents féminins par rapport aux agents masculins. Aussi, ils ont trouvé les agents féminins comme étant plus utiles, plus efficaces et mieux conçus, plus captivants, plus engageants et plus attrayants que les agents masculins (ESPOSITO et al., 2021). Cependant, dans une expérience similaire avec des agents silencieux, le genre des agents n'affectait pas les participants selon les mêmes critères (ESPOSITO et al., 2021). Les voix pourraient avoir influencé le niveau de masculinité/féminité perçu des agents, mais cela n'a pas été mesuré dans ces études. Dans une autre étude, les agents experts ont été jugés plus crédibles que les agents motivateurs, quel que soit leur genre (KANTHARAJU et al., 2018).

Vers une évolution des perceptions stéréotypées? Une étude récente de NAG et YALÇIN (2020) contredit les recherches antérieures sur la façon dont les individus perçoivent les agents virtuels en fonction de leur genre : les images fixes d'agents masculins et féminins ont été évaluées de la même manière pour les traits d'*agentivité* (traits généralement associés aux hommes : *ambitieux, agressif, courageux, décisif*) et de *compétence* (traits généralement associés aux hommes : *créatif, intelligent, innovant, organisé*). En revanche, en cohérence avec les autres études, ce n'était pas le cas pour les traits de *communalité* (traits généralement associés aux femmes : *affectueux, compatissant, sensible, invitant, serviable*) où les agents féminins étaient mieux notés. Une limite de cette étude est que les agents féminins et masculins étaient assez similaires en apparence. Ceci étant dit, les résultats de l'étude tendent à être cohérents avec l'évolution des stéréotypes de genre rapportés par EAGLY et al. (2020) pour la perception des traits d'*agentivité* et de *compétence* dans les interactions interpersonnelles : l'écart de ces traits en faveur des hommes a diminué, et les traits de *communalité* sont encore largement attribués aux femmes. Cela pose la question de savoir si l'évolution de la perception des stéréotypes dans les interactions humain-humain montrée par EAGLY et al. (2020) peut être observée de manière similaire dans les interactions humain-agent.

Discussion. Sur la base des recherches présentées ci-dessus, il semble que les agents virtuels masculins soient perçus comme plus compétents, en particulier sur les sujets

stéréotypés liés aux hommes où les individus tendent à suivre leurs conseils par rapport aux agents virtuels féminins. Ils semblent mieux adaptés pour représenter un agent virtuel pédagogique dans les domaines des STIMs, puisque ces domaines sont perçus comme masculins. Dans la sous-section suivante, nous nous intéressons plus particulièrement aux recherches sur les agents virtuels pédagogiques et l'impact de leur genre sur les résultats scolaires des utilisateurs et utilisatrices.

3.2.4. Effet du genre des agents virtuels pédagogiques sur les résultats scolaires

Nous nous intéressons ici à l'effet du genre des agents virtuels pédagogiques, plus spécifiquement dans un contexte d'apprentissage. Nous avons résumé les études sélectionnées sur l'impact du genre des agents virtuels pédagogiques dans le cadre d'une tâche d'apprentissage dans chaque ligne du Tableau 3.2. Nous avons indiqué les caractéristiques des agents virtuels, le nombre de participants hommes et femmes avec leur âge moyen, les tâches de l'étude, les mesures observées et les résultats de l'étude.

TABLEAU 3.2

Synthèse d'articles sur l'impact des agents virtuels genrés dans le cadre d'une tâche d'apprentissage. Les articles sont classés du plus ancien au plus récent. AF = Agent(s) Femme(s), AH = Agent(s) Homme(s), PF = Participante(s) Femme(s), PH = Participant(s) Homme(s). La colonne des agents décrit le nombre d'agents en fonction de leur genre, leur dimension (2D ou 3D), leur apparence (réaliste ou cartoon) et leur rôle. La colonne des participants décrit le nombre d'hommes et de femmes qui ont participé à l'étude, ainsi que l'âge moyen. Dans la colonne des résultats, par exemple, "PH > PF" signifie que cela a eu plus d'impact sur les participants hommes que sur les participantes. "AF > AH" signifie que l'agent femme a eu plus d'impact que l'agent homme. Les explications se trouvent après ce tableau.

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
K. N. MORENO et al. (2002)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Tuteur	• 12 PH • 27 PF • avg age 20	Regarder une vidéo d'un agent virtuel donnant un cours, passer un test à choix multiples.	• Performance • Masculinité, féminité perçu	• *Performance* : AH > AF • *Masculinité, féminité perçu* : - AF : très féminin - AH : masculin
BAYLOR et KIM (2004)	• 4 AH • 4 AF • 2D, 3D • Voix • Réaliste, cartoon • Tuteur	• 94 PH • 218 PF • avg age 20.54	Créer un emploi du temps avec l'aide d'un agent virtuel.	• Auto-efficacité • Autorégulation • Connaissance • Intelligence • Apprentissage	• *Auto-efficacité, autorégulation, connaissance, intelligence* : AH > AF • *Apprentissage* : pas significatif

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
BAYLOR et KIM (2004)	• 6 AH • 6 AF • 2D, 3D • Voix • Réaliste, cartoon • Expert, motivateur, mentor	• 89 PH • 140 PF • avg age 19.39	Créer un emploi du temps avec l'aide d'un agent virtuel.	• Connaissance • Intelligence • Apprentissage • Autorégulation • Auto-efficacité	• *Connaissance, intelligence* : AH > AF • *Apprentissage, autorégulation :* pas significatif • *Auto-efficacité* : AF > AH
R. MORENO et FLOWERDAY (2006)	• 5 AH • 5 AF • 2D • Voix • Réaliste • Tuteur	• 21 PH • 59 PF • avg age 26.88	Regarder une vidéo d'un cours dispensé par un agent virtuel, passer un test.	• Aide • Motivation • Agent choisi • Apprentissage	• *Aide, motivation, apprentissage* : pas significatif • *Agent choisi* : agent de même genre = agent de genre opposé
Y. KIM et al. (2007)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Compagnon	• 11 PH • 45 PF • avg age 20.71	Créer un cours sur les concepts économiques avec l'aide d'un agent virtuel.	• Faciliter l'apprentissage • Engageant • Humain • Apprentissage (rappel)	• *Faciliter l'apprentissage, engageant, humain* : AH > AF • *Apprentissage (rappel)* : pas significatif

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
PLANT et al. (2009)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Orateur	• 45 PH • 61 PF • avg age 13.63	Écouter une histoire sur 4 ingénieures et les avantages de l'ingénierie, dispensée soit par un agent, soit par voix uniquement. Passer un test de mathématiques.	• Intérêt • Utilité • Auto-efficacité • Performance • Moins de stéréotypes de genre en ingénierie	• *Intérêt, utilité* : AF > AH & no agent • *Auto-efficacité* : AH & AF > sans agent • *Performance* : AF > AH • *Moins de stéréotypes de genre en ingénierie* : - PH : agents > sans agent - PF : AF & sans agent > AH
HAYES et al. (2010)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Observateur	• 35 PH • avg age 19.77	Contrôler un avatar lors d'un test de mathématiques (vue à la 1ère ou 3ème personne), en présence d'un agent masculin ou féminin, ou sans agent.	• Présence sociale • Performance • Temps de réponse	• *Présence sociale* : AH > AF & sans agent • *Performance, temps de réponse* : - 1ère personne : sans agent & AH > AF - 3ème personne : AF > sans agent & AH
Y. KIM et WEI (2011)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Tuteur	• 110 PH • 100 PF • avg age 15.93	Passer un test de mathématiques sans agent, regarder un agent expliquer la leçon, résoudre des problèmes de mathématiques avec l'agent (entraînement), passer un 2ème test de mathématiques sans agent.	• Agent choisi • Performance	• *Agent choisi* : agents de même genre & agents de même origine culturelle > autres • *Performance* : tout le monde s'est amélioré

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
SILVERVARG et al. (2013)	• 1 AH • 1 AF • 1 neutre • 2D • Texte • Cartoon • Élève	• 46 PH • 37 PF • 12-14 ans	Interagir avec un élève androgyne virtuel sur un cours de mathématiques, puis avec un élève virtuel (féminin ou masculin).	• Androgynie perçue • Agent préféré comme élève • Agent préféré comme partenaire de discussion	• *Androgynie perçue* : agent androgyne = androgyne • *Agent préféré comme élève* : - PF : agent androgyne > AH & AF - PH : agent androgyne > AH & AF (pas significatif) • *Agent préféré comme partenaire de discussion* : - PF : agent androgyne > AH - PH : agent androgyne > AF
Y. KIM et LIM (2013)	• 2 AF • 3D • Voix • Réaliste • Tuteur	• 64 PH • 56 PF • avg age 15.93	Passer un test de mathématiques sans agent, regarder un agent expliquer la leçon, résoudre des problèmes de mathématiques avec l'agent (entraînement), passer un 2ème test de mathématiques sans agent.	• Performance • Auto-efficacité	• *Performance* : tout le monde s'est amélioré • *Auto-efficacité* : - PF : agent présent > sans agent - PH : pas d'amélioration
Y. KIM (2013)	• 1 AH • 1 AF • 1 neutre •2D, 3D • Voix, texte • Cartoon • Tuteur	• 68 PH • 73 PF • avg age N/A	Répondre à des questions sur un texte posées par un agent virtuel.	• Performance	• *Performance* : - PF = PH - PF : AH & AF > agent robotique - PH : AH > AF & agent robotique

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
A. M. JOHNSON et al. (2013)	• 1 AH • 1 AF • 3D • Voix • Cartoon • Tuteur	• 88 PH • 109 PF • avg age 12.1	Regarder un agent enseigner une leçon sur les circuits électriques, passer un test à choix multiples.	• Performance • Évaluation du programme	• *Performance* : pas significatif • *Évaluation du programme* : AF > AH
OZOGUL et al. (2013)	• 2 AH • 2 AF • 3D • Voix • Cartoon • Tuteur	• 173 PH • 161 PF • avg age 12.3	Regarder un agent (choisi ou assigné au hasard) enseigner une leçon sur les circuits électriques, passer un test à choix multiples.	• Performance • Évaluation du programme • Agent choisi	• *Performance* : - Agent assigné : pas significatif - Agent choisi : agent de genre opposé > agent de même genre • *Évaluation du programme* : pas significatif • *Agent choisi* : agent de même genre > agent de genre opposé
SHIBAN et al. (2015)	• 1 AH • 1 AF • 3D • Texte • Cartoon • Tuteur	• 21 PH • 73 PF • avg age 20.20	Passer un test de mathématiques pendant qu'un agent virtuel donnait des feedbacks.	• Intérêt • Motivation • Plaisir • Crédible • Engageant • Humain • Faciliter l'apprentissage • Performance	• *Intérêt, motivation* : AF > AH • *Plaisir* : pas significatif • *Crédible, engageant, humain* : AH > AF • *Faciliter l'apprentissage* : pas significatif • *Performance* : AH > AF (légèrement)

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
Y. KIM (2016)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Tuteur	• 67 PF • avg age 15.51	Écouter un agent parler de manière convaincante des avantages des STIMs, résoudre des problèmes de mathématiques avec le même agent, résoudre des problèmes de mathématiques sans l'agent.	• Crédibilité • Convivialité • Aide • Attitudes positives pour apprendre les mathématiques	• *Crédibilité, convivialité, aide* : - Participants issus de minorités culturelles : - AH : agent pair > agent tuteur - AF : agent tuteur > agent pair - Participants caucasiens : - AH & AF : pas significatif • *Attitudes positives pour apprendre les mathématiques* : - Participants issus de minorités culturelles : - AH : agent pair > agent tuteur - AF : agent tuteur > agent pair - Participants caucasiens : - AH & AF : pas significatif
KRÄMER et al. (2016)	• 2 AH • 2 AF • 3D • Voix • Réaliste • Interviewer motivateur	• 60 PH • 68 PF • avg age 23.85	Passer un test de mathématiques sans agent, puis passer un test de mathématiques avec un agent présent expliquant la procédure.	• Motivation • Sentiment de rapport • Performance	• *Motivation, sentiment de rapport* : pas significatif • *Performance* : - PF & agent-rapport : AH > AF - PH & agent-rapport : AF > AH
LI et al. (2016)	• 1 AH • 1 neutre • 3D • Voix • Réaliste • Tuteur	• 20 PH • 20 PF • avg age 20.48	Regarder un agent présenter des diapositives sur des cours sur l'interaction humain-machine.	• Apprentissage	• *Apprentissage* : - PH : agent robot > vrai homme > AH > image fixe d'un robot - PF : pas de différences

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
JEONG et al. (2017)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Instructeur	• 54 PH • 63 PF • avg age 20.94	Écouter les feedbacks négatifs d'un agent instructeur tout en jouant une scène. Reproduire la scène avec l'agent instructeur et un agent élève (pas de feedback).	• Avancer • Reculer	• *Avancer* : - PF : AF > AH - PH : AH > AF • *Reculer* : - PF : AH > AF - PH : AF > AH
PEZZULLO et al. (2017)	• 1 AF • 3D • Voix • Réaliste • Compagnon	• 54 PH • 63 PF • avg age 13.30	Jouer à un jeu sur les cours de biologie avec l'aide d'un agent virtuel.	• Demande mentale • Engagement avec l'agent • Performance	• *Demande mentale, engagement avec l'agent* : PF > PH • *Performance* : PF = PH
WIRZBERGER et al. (2019)	• 1 AH • 3D • Voix • Réaliste • Instructeur	• 27 PH • 35 PF • avg age 69.03	Mémoriser une liste de mots après avoir suivi un entraînement de mémorisation par un agent.	• Apprentissage (rappel)	• *Apprentissage (rappel)* : PF > PH
MAKRANSKY et al. (2019)	• 1 AF • 1 neutre • 3D • Voix • Réaliste • Tuteur	• 33 PH • 33 PF • avg age N/A	Regarder un agent virtuel enseigner la sécurité en laboratoire, passer des tests.	• Présence sociale • Apprentissage (rappel et transfert de connaissance)	• *Présence sociale* : - PF : AF = agent drone - PH : AF > agent drone • *Apprentissage* : - PF = PH - PF : AF > agent drone - PH : agent drone > AF

Référence	Agent(s)	Participant(s)	Tâche(s)	Mesure(s)	Résultat(s)
CHANG et al. (2019)	• 1 AH • 3D • Voix • Réaliste • Instructeur	• 76 PF • avg age N/A	Contrôler un avatar masculin ou féminin, apprendre à résoudre des problèmes arithmétiques en présence d'un agent masculin (dominant ou non dominant, en fonction de sa posture corporelle), résoudre des problèmes sans la présence de l'agent.	• Apprentissage (rappel & performance)	• *Apprentissage* : - agent non-dominant > agent dominant - pas d'effet du genre de l'avatar
SAJJADI et al. (2020)	• 1 AH • 1 AF • 3D • Voix • Réaliste • Instructeur	• 8 PH • 4 PF • avg age 19.6	Observer des formations géologiques dans un environnement virtuel, répondre aux questions posées par un agent.	• Efficacité d'apprentissage perçu • Apprentissage • Leadership • Convivialité • Présence sociale et spatiale	• *Efficacité d'apprentissage perçu* : AF > AH • *Apprentissage* : pas significatif • *Leadership, convivialité, présence sociale et spatiale* : AF > AH (pas significatif)
SPILIOTOPOULOS et al. (2020)	• 1 AF • 3D • Voix • Réaliste • Tuteur	• 24 PH • 16 PF • avg age 20	Apprendre grâce à un agent virtuel comment argumenter, comment être empathique aux besoins des autres, comment négocier.	• Auto-efficacité • Facilité du système • Aide • Apprentissage	• *Auto-efficacité* : pas significatif • *Facilité du système* : PF > PH • *Aide* : PH > PF • *Apprentissage* : pas significatif (tout le monde s'est amélioré)

Effets positifs des agents féminins. Diverses études sur les environnements d'apprentissage virtuels (Tableau 3.2) ont rapporté que le genre d'un agent virtuel pédagogique peut avoir un impact sur les performances d'apprentissage des utilisateurs. Dans une récente étude, MAKRANSKY et al. (2019) a montré que les jeunes filles participantes réussissaient mieux des exercices de sécurité de laboratoire scientifique (en termes d'apprentissage et d'apprentissage par transfert) lorsque l'agent virtuel pédagogique était une scientifique virtuelle plutôt qu'un drone virtuel. L'inverse était vrai pour les garçons. Les chercheurs ont suggéré que les garçons s'identifiaient au drone, tandis que les filles s'identifiaient à l'agent virtuel pédagogique féminin. Une limite de cette étude est qu'il n'y a pas de comparaison avec un agent masculin, ni de mesure d'identification aux agents virtuels pédagogiques. Dans une étude de SHIBAN et al. (2015), les apprenantes étaient plus motivées et intéressées par les mathématiques lorsqu'elles étaient enseignées par un agent féminin par rapport à un agent masculin. Cependant, elles ont obtenu de meilleurs résultats avec l'agent masculin, ce qui peut s'expliquer par leur perception de l'apparence des agents virtuels pédagogiques. L'agent masculin était plus âgé et portait une cravate, tandis que l'agent féminin était jeune et joli. Selon les auteurs, les performances des participantes se sont améliorées car l'agent masculin était perçu comme un expert, notamment à cause de son apparence et de son âge. Il a en effet été montré dans d'autres études que les agents virtuels perçus comme des experts amélioraient les performances des apprenants (BAYLOR & KIM, 2004). Les chercheurs ont également conclu que la motivation et l'intérêt des participantes s'amélioraient avec l'agent féminin en raison de leur similitude avec l'agent (en âge et en genre), conformément à "l'hypothèse de similitude" et en cohérence avec les résultats d'autres études (R. ROSENBERG-KIMA et al., 2008). Cet argument est cohérent avec la théorie de l'apprentissage cognitif social de BANDURA (1986) : les gens apprennent souvent en imitant des personnes qu'ils perçoivent comme similaires (ou supérieures en rang ou en statut) à eux et qui sont donc acceptées comme modèles sociaux (Section 2.3.4.3). Cette théorie est également confirmée dans une étude de PLANT et al. (2009), où un agent féminin a augmenté le sentiment de réussite des participants, filles et garçons, en délivrant un message sur les avantages de l'ingénierie, ce qui a entraîné de meilleures performances et un plus grand intérêt pour les mathématiques.

Effets positifs des agents masculins. D'autres recherches ont démontré un effet positif des agents masculins par rapport aux agents féminins dans des tâches pédagogiques. Par exemple, dans deux expériences menées par BAYLOR et KIM (2004), un agent virtuel a aidé les participants à créer un emploi du temps. Le genre de l'agent n'a pas eu d'impact sur l'apprentissage mais a affecté le sentiment de réussite, qui a augmenté davantage dans la première expérience avec l'agent masculin que l'agent féminin ; le contraire s'est produit dans la deuxième expérience pour les participants, hommes comme femmes. Les chercheuses ont expliqué les résultats de la première expérience par la perception des agents masculins : les participants ont évalué l'agent masculin comme étant plus intéressant et utile que l'agent féminin. Dans la deuxième expérience, les participants considéraient l'agent féminin comme moins intelligent

que l'agent masculin, bien qu'ils aient reçu les mêmes instructions de la part des deux agents. Certaines recherches ont indiqué que les agents perçus comme moins intelligents pourraient conduire à un plus grand sentiment de réussite, en partant du principe que « si même lui peut le faire, alors moi aussi » (BAYLOR & KIM, 2005). Dans une expérience similaire menée par Y. KIM et al. (2007), les chercheuses ont introduit un agent virtuel pédagogique féminin et un autre agent virtuel pédagogique masculin pour aider les étudiants à concevoir un cours d'apprentissage en ligne, qui comprenait la création d'un planning. Les étudiants travaillant avec l'agent masculin l'ont mieux noté pour ce qui est de faciliter l'apprentissage, d'être engageant et d'être humain que les étudiants travaillant avec l'agent féminin. Notamment, l'agent masculin a eu un impact plus positif que l'agent féminin sur l'intérêt et l'apprentissage des participants en termes de rappel (la capacité de se souvenir de ce que l'agent a dit pendant la tâche).

Stéréotypes sur le niveau de féminité. D'autres facteurs peuvent également entrer en jeu dans les études sur l'impact des agents virtuels sur l'apprentissage. Dans l'étude de K. N. MORENO et al. (2002), les participants ont visionné une vidéo d'un agent virtuel présentant un cours sur la tension artérielle, suivi d'un test à choix multiples. Les résultats de cette étude suggèrent que les participants ont plus appris avec l'agent masculin que l'agent féminin. Les chercheurs ont suggéré que cela pourrait être dû au fait que l'agent féminin ne se conformait pas au stéréotype du professeur, généralement masculin. L'étude pilote de cette recherche a montré que les participants percevaient l'agent féminin comme très féminin, et l'agent masculin comme très masculin. Les participants ont peut-être jugé les agents féminins comme étant "trop féminin" pour convenir au rôle de professeur. L'étude n'a cependant pas abordé la façon dont les participants percevaient l'expertise des agents ni cherché à déterminer les éventuelles interactions entre l'expertise perçue, la féminité perçue de l'agent et la performance au test.

Discussion. Les recherches présentées ci-dessus mettent en évidence l'importance du genre des agents virtuels pédagogiques sur les apprentissages. Différentes études semblent donner des résultats contradictoires, d'une part en montrant que l'apprentissage s'améliore lorsque le genre des agents virtuels correspond à celui de l'apprenant, mais d'autre part que les agents virtuels masculins pourraient être mieux adaptés à l'amélioration de l'apprentissage. La Section 3.2.3 montre que les agents virtuels masculins sont perçus comme plus compétents que les agents féminins, et que les utilisateurs suivent en général plus les conseils d'un agent masculin que d'un agent féminin sur des sujets perçus comme masculins. Cependant, certaines études mettant en scène un agent féminin dans les domaines des STIMs présentées dans cette section montrent que ces agents féminins ont une influence positive sur les résultats scolaires : ils améliorent l'apprentissage, le sentiment de réussite, l'intérêt et la motivation, malgré le fait que les STIMs soient perçus comme des domaines masculins, et ce en particulier pour les filles et les femmes.

En conclusion, à ce jour les études présentées dans cette section ne permettent pas d'affirmer qu'un genre est plus adapté qu'un autre pour l'apprentissage. De plus, pra-

tiquement toutes ces études ont été réalisées avec des participants adultes et non avec des enfants. Il est cependant à noter que chaque étude a été réalisée dans un contexte d'apprentissage spécifique avec une pédagogie, une conception de l'environnement virtuel, une durée de l'interaction, une modalité d'interaction, ou encore un environnement physique spécifique. Ces éléments de contexte peuvent avoir un impact sur les perceptions des utilisateurs et les résultats d'apprentissage. Une limite de cette revue est que nous n'avons pas pris en compte toutes ces spécificités contextuelles. Une analyse plus approfondie pourrait prendre en compte ces éléments contextuels pour fournir une vision plus globale de l'impact du genre des agents virtuels sur les apprentissages scolaires. Aussi, peu d'études présentées dans cette section utilisait des agents virtuels en tant que modèle de réussite. Dans cette thèse, notre objectif est d'explorer l'influence du genre d'agents virtuels pédagogiques sur les performances des enfants en mathématiques, lorsque ces agents incarnent un modèle de réussite. Peu de travaux ont étudié la simulation de modèles de réussite à travers des agents virtuels pédagogiques. Nous les présentons dans la section suivante.

3.3. Agents virtuels pédagogiques en tant que modèles de réussite

En cognition sociale, les filles et les femmes, en tant que modèles de réussite compétents dans leur domaine, permettent de réduire les effets de menace du stéréotype (Section 2.3.4.3). Le mécanisme d'identification à ce modèle de réussite social (e.g., s'inspirer de la personne et vouloir lui ressembler) permet aux apprenants de réduire leur anxiété, d'améliorer leurs résultats, et de se dire qu'eux aussi, ils peuvent y arriver. Parmi les attributs recherchés chez ces modèles de réussite, on retrouve (en plus des compétences) leur statut, leur âge, et des similitudes (genre, origine culturelle, loisirs, ...). Les apprenants tendent à se rattacher à des individus un peu plus âgés ayant un statut plus élevé. De plus, la réussite du modèle de réussite ne doit pas être perçue comme inatteignable. Ainsi, un agent virtuel pédagogique légèrement plus âgé que les utilisateurs et similaire à eux, présentant une réussite atteignable dans son domaine, aurait à priori les caractéristiques nécessaires pour être un bon modèle de réussite. Dans cette section, nous nous intéressons aux agents virtuels pédagogiques genrés, incarnant des modèles de réussite, plus particulièrement en mathématiques et en sciences.

Rôles pédagogiques. À notre connaissance, il y a peu de travaux sur les modèles de réussite incarnés par des agents virtuels en mathématiques. On peut tout de même rapprocher le modèle de réussite au rôle appelé *mentor* par BAYLOR et KIM (2005) et Y. KIM et BAYLOR (2016). Cet agent présente plusieurs similitudes avec le modèle de réussite tel que défini en sciences sociales. Un agent mentor combine des caractéristiques d'un agent expert grâce à son expertise accrue, tout en incorporant des éléments d'un agent motivateur par son enthousiasme et ses encouragements. Tout comme un modèle de réussite, un agent mentor est compétent dans son domaine

et légèrement plus âgé que les apprenants. Il exprime de plus des comportements positifs de sorte à créer une interaction plus positive. Les différents rôles pédagogiques tels qu'étudiés par BAYLOR et KIM (2005) sont illustrés en Figure 3.1.

FIGURE 3.1.

*Caractéristiques des agents expert, motivateur, et mentor (*BAYLOR & KIM, *2005).*

	Expert	**Motivator**	**Mentor**
Image			
Animation	*Deictic*	*Emotional*	*Deictic & Emotional*
Voice	Authoritative & Monotone	Effusive & Enthusiastic	Confident & Calm
Affect	None	▪ Acknowledgment ▪ Confusion ▪ Disapproval ▪ Excitement ▪ Pleasure ▪ Surprise	▪ Acknowledgment ▪ Confusion ▪ Disapproval ▪ Excitement ▪ Pleasure ▪ Surprise
Script	Information	Encouragement	Information & Encouragement

Plusieurs travaux ont évalué l'impact des agents virtuels pédagogiques suivant leur rôle. Par exemple, dans les études de BAYLOR et KIM (2005) et Y. KIM et BAYLOR (2016), l'agent mentor a permis aux participants d'augmenter leurs performances et leur motivation dans une tâche de création de cours de mathématiques. Cela pourrait s'expliquer par le fait que l'agent mentor est celui qui se rapproche le plus d'un modèle de réussite. Comme expliqué Section 3.2.4, dans l'étude de SHIBAN et al. (2015), il se pourrait que l'agent virtuel pédagogique masculin ait été perçu comme un agent expert. En effet, cet agent était âgé et portait une cravate, et il a permis aux apprenantes d'améliorer leur performance en mathématiques. L'agent virtuel pédagogique féminin, quant à lui, a permis aux apprenantes d'augmenter leur motivation et leur intérêt en mathématiques. Cet agent aurait pu être perçu comme un modèle de réussite ; cependant, l'agent était également d'apparence très féminine (voir Figure 3.2 pour les apparences des agents). Comme expliqué en Section 2.3.4.3, dans le domaine des STIMs, la féminité des modèles de réussite peut entraver l'effet habituellement positif de ces modèles sur l'apprentissage. Ainsi, dans cette thèse, nos agents virtuels

masculin et féminin portent les mêmes vêtements pour éviter le plus possible qu'il y ait une différence de perception selon leurs vêtements (Section 6.2.4).

FIGURE 3.2.

Agents masculin et féminin de l'étude de SHIBAN *et al. (2015).*



Modèles de réussite féminins. Certaines recherches, bien qu'encore très limitées, ont exploré l'utilisation d'agents virtuels pédagogiques incarnant des modèles de réussite pour augmenter les performances et l'intérêt des apprenants en mathématiques. R. ROSENBERG-KIMA et al. (2008) ont montré l'efficacité d'une ingénieure virtuelle pour intéresser les femmes dans les domaines des STIMs. Dans une vidéo, l'agent, qui était du même genre que les participantes (toutes des femmes), a présenté l'histoire de plusieurs femmes réussissant dans les domaines des STIMs tout en se présentant comme étant ingénieure. Cela a conduit à un changement d'attitude des participantes envers les STIMs, comme l'a montré un questionnaire sur une échelle de 7 points. Les femmes ayant vu un agent virtuel féminin étaient moins susceptibles d'approuver les stéréotypes de genre traditionnels en STIMs que celles ayant vu un agent virtuel masculin. Elles étaient également plus susceptibles de croire que les femmes pouvaient réussir dans ces domaines. Cependant, quelques stéréotypes de genre persistaient : les participantes étaient légèrement plus susceptibles de croire en l'utilité des STIMs avec un ingénieur virtuel masculin. Une autre étude similaire de R. B. ROSENBERG-KIMA et al. (2010) a montré que les agents virtuels Noirs avaient un impact plus positif sur l'intérêt pour les STIMs et les stéréotypes de genre en STIMs pour les femmes Noires, tandis que les agents virtuels féminins (Noirs ou Blancs) avaient un impact plus positif sur les femmes Blanches sur les mêmes critères. Ces recherches montrent l'importance des ressemblances physiques entre les humains et les agents, telle que l'origine culturelle et le genre, sur la performance et l'intérêt pour les mathématiques.

Modèles de réussite avec les enfants. Ces études se concentraient cependant sur des participants adultes. PLANT et al. (2009) ont mené une étude similaire, cette fois-ci avec des participants enfants entre 12 et 15 ans. Les filles et les garçons ont obtenu de meilleurs résultats et étaient plus intéressés par l'ingénierie après avoir interagi avec un agent féminin : cet agent a augmenté leur sentiment de réussite et leurs évaluations

sur l'utilité des STIMs. Les participants garçons étaient moins susceptibles d'approuver les stéréotypes de genre traditionnels en STIMs en présence d'un agent, masculin ou féminin. Les participantes, quant à elles, étaient moins susceptibles d'approuver ce genre de stéréotype avec un agent féminin ou sans aucun agent, plutôt qu'avec un agent masculin. Cependant, les agents virtuels de cette étude étaient des adultes, alors qu'un bon modèle social doit être légèrement plus âgé que les participants pour permettre l'identification. Y. KIM (2013) a étudié l'effet d'agents virtuels pédagogiques enfants, fille et garçon, ainsi que d'un agent virtuel pédagogique ressemblant à un robot sur les performances en lecture. Les agents humanoïdes ressemblaient aux participants, enfants également, et pouvaient être perçus comme des modèles de réussite. La compréhension de texte des filles participantes a augmenté avec un agent humanoïde (fille et garçon), alors que les garçons ont mieux compris le texte uniquement avec un agent masculin. Il se pourrait que l'agent masculin ait été efficace, en tant que modèle de réussite, sur les effets de menace du stéréotype en lecture chez les garçons (PANSU et al., 2016).

D'autres recherches ont également étudié les pairs virtuels, similaires à des modèles de réussite, en se focalisant sur d'autres caractéristiques que le genre. Par exemple, IACOBELLI et CASSELL (2007) ont développé des agents virtuels enfants modélisant l'origine culturelle à travers le comportement verbal et non-verbal. FINKELSTEIN et al. (2013) ont alors montré l'effet positif d'un agent virtuel pédagogique Noir et enfant, parlant en anglais afro-américain, sur l'apprentissage en sciences. Les participants enfants, tous Afros-américains, se sont plus améliorés en matière de discours scientifique (formuler des hypothèses cohérentes) avec un tel agent, comparé au même agent virtuel pédagogique parlant en anglais-américain standard. De plus, l'agent parlant en anglais afro-américain n'a pas changé le dialecte employé par les enfants lors de la présentation de leur discours scientifique. Les autrices ont précisé que cela a des implications importantes, car les enseignants craignent que le fait d'autoriser le dialecte afro-américain dans leur classe rende les élèves moins susceptibles d'utiliser le dialecte standard. Dans cette thèse, nous avons suivi cette recherche pour modéliser le lexique de nos agents virtuels pédagogiques afin qu'il se rapproche le plus possible de celui des enfants participants.

En résumé, les agents virtuels pédagogiques pourraient incarner des modèles de réussite à travers le rôle (comme l'agent mentor) et l'apparence (un peu plus âgé, genre, origine culturelle, vêtements non-strictes). Une limite que l'on remarque cependant dans ces études est qu'il n'y a pas de mesure d'identification à l'agent virtuel. Nous ne pouvons savoir à quel point les participants ont perçu les agents comme des modèles de réussite. De plus, il y a peu de travaux avec des enfants participants. Dans cette thèse, nous étudions la modélisation d'un modèle de réussite à travers le rôle de mentor tel que décrit dans cette section, l'apparence, et des mesures d'identification à l'agent virtuel. D'autres capacités peuvent également être importantes pour un modèle de réussite, comme l'établissement d'une interaction positive et coordonnée avec l'apprenant pour percevoir plus de similitudes. Nous présentons ce type d'interaction dans la prochaine section.

3.4. Le rapport virtuel pour améliorer les performances académiques

Le rapport correspond au sentiment de connexion et d'harmonie lors d'une interaction positive (BERNIERI, 1988 ; TICKLE-DEGNEN & ROSENTHAL, 1990), atteignable à l'aide de trois composantes : la positivité, l'attention mutuelle, et la coordination. Le rapport *virtuel* a été implémenté lors d'interaction entre agent virtuel et humain. En effet, certains comportements verbaux et non-verbaux favorisent le rapport, alors que d'autres comportement, comme la dominance, tendent à le diminuer (R. ZHAO et al., 2014). Comme montré dans l'étude de CHANG et al. (2019) en réalité virtuelle, un agent pédagogique masculin avec des comportements perçus comme dominants (e.g., s'asseoir proche des participantes avec une posture ouverte) ont baissé les performances en arithmétiques des participantes, comparé à un agent pédagogique masculin perçu comme non-dominant (e.g. s'asseoir plus loin avec une posture plus fermée). De plus, KRÄMER et al. (2016) ont analysé l'impact du genre des agents virtuels couplé avec leur comportement verbal et non-verbal sur la motivation, l'effort, et la performance en mathématiques de participants adultes. Ils ont montré que la simple présence d'un agent virtuel féminin n'a pas augmenté la motivation des participantes. En revanche, quand l'agent a démontré un comportement non-verbal s'alignant sur celui des participants (GRATCH et al., 2007), la performance et l'effort des participants ont augmenté. Cette étude a alors montré l'importance du **rapport virtuel** dans le domaine de l'éducation.

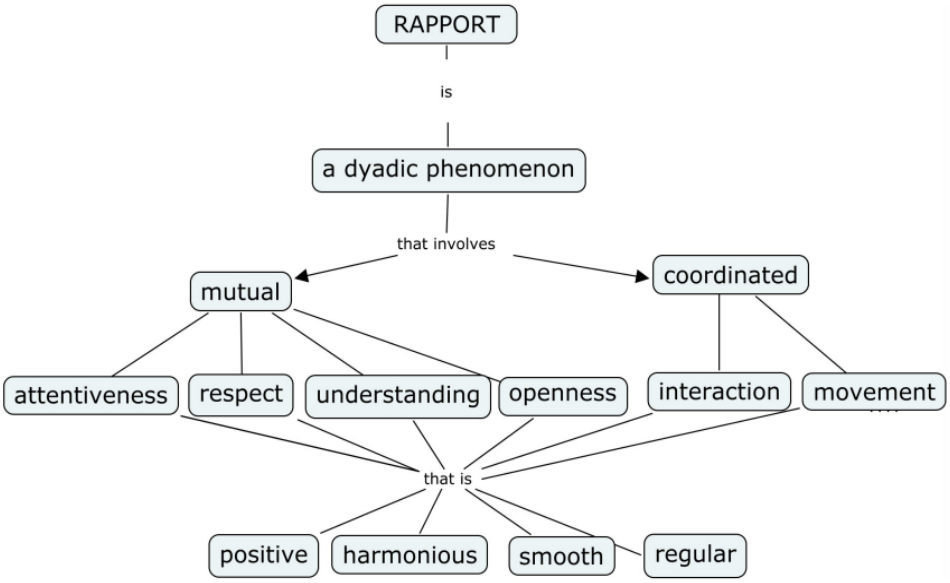
Dans cette section, nous introduisons tout d'abord une définition théorique du rapport. Nous présentons ensuite des études modélisant le rapport virtuel dans les interactions entre un agent virtuel et un utilisateur.

3.4.1. Fondement théorique du rapport

Les communications interpersonnelles étudient l'interaction sociale entre les individus, pour comprendre quels comportements verbaux et non-verbaux sont utilisés pour mettre en œuvre une stratégie de communication, que ce soit pour informer, persuader, ou fournir un soutien émotionnel aux autres (HARTLEY, 2002). Nous visons ici à comprendre comment les individus établissent une interaction positive via le rapport. Une définition généralisée du rapport est le sentiment de connexion et d'harmonie ressentie lors d'une interaction. De manière informelle, le rapport dans une interaction dyadique est souvent reconnu par les individus lorsqu'ils se sont sentis "connectés" les uns aux autres et qu'ils ont perçu l'interaction comme agréable (BERNIERI, 1988 ; TICKLE-DEGNEN & ROSENTHAL, 1990). La Figure 3.3 schématise les éléments du rapport pour parvenir à établir une interaction harmonieuse (E. MURPHY & RODRÍGUEZ-MANZANARES, 2012).

FIGURE 3.3.

Éléments du rapport (E. MURPHY & RODRÍGUEZ-MANZANARES, 2012).



De nombreuses études ont montré que le rapport influence positivement la communication dans divers cas et domaines (GREMLER & GWINNER, 2000 ; S. KIM et al., 2020 ; MACINTOSH, 2009). Dans les interactions en classe, le rapport est un facteur important pour les résultats scolaires. Les comportements des enseignants peuvent avoir un impact différent sur l'apprentissage des filles et des garçons : l'enquête à grande échelle de FORGASZ et LEDER (1996) a montré que les élèves qui percevaient que leurs professeurs de mathématiques s'intéressaient à eux en tant qu'individus étaient plus susceptibles de croire en leur capacité en mathématiques, et cela était d'autant plus vrai pour les filles que pour les garçons. Pas seulement en cours de mathématiques, l'établissement d'une bonne relation entre l'enseignant et les élèves en général est extrêmement important pour créer un climat de classe confortable où les élèves peuvent mieux apprendre (TSUI, 1996). BENSON et al. (2005) ont constaté dans leur sondage que dans les classes où les enseignants établissaient le rapport, les étudiants étaient plus susceptibles d'assister aux cours, d'être attentifs, et d'apprécier le sujet du cours. Aussi, les étudiants participaient plus en classe et avaient de meilleurs résultats scolaires (FRISBY & MARTIN, 2010 ; GRANITZ et al., 2009).

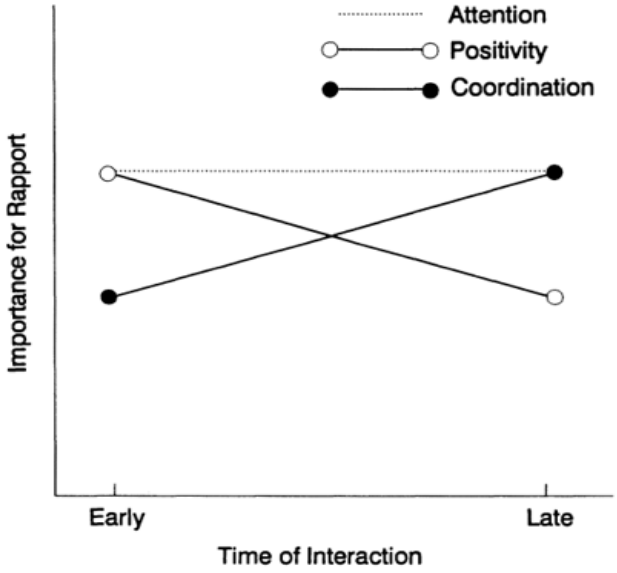
Le rapport peut être manifesté à travers le comportement verbal, mais également à l'aide du comportement non-verbal via les gestes et les expressions faciales. Selon TICKLE-DEGNEN et ROSENTHAL (1990), le rapport entre les interactants peut-être atteint à l'aide de 3 composantes : (1) une attitude positive (e.g., hochement de tête), (2) une attention mutuelle (e.g., ne pas regarder ailleurs), et (3) une coordination du comportement verbal et non-verbal (e.g., synchronie, mimicrie). Les mêmes auteurs ont défini les postures du corps comme exprimant de l'attention, les expressions faciales comme exprimant de la positivité, et les comportements de mimicrie et de

synchronie comme exprimant de la coordination. On peut adapter ses mouvements du corps ainsi que le ton de sa voix à son interlocuteur (BERNIERI, 1988; NICKELS et al., 1983), ou encore partager les mêmes gestes et postures par mimétisme (GILLIS et al., 1995; GRANITZ et al., 2009; LAKIN & CHARTRAND, 2003). Du côté verbal, on peut employer des stratégies de communication (e.g., faire des révélations sur soi, féliciter) ou encore utiliser le même lexique que son interlocuteur pour permettre la coordination (CASSELL & TVERSKY, 2005).

Bien que les 3 composantes du rapport soient liées, TICKLE-DEGNEN et ROSENTHAL (1990) soutiennent qu'elles ne sont pas équivalentes. Par exemple, une interaction peut être positive mais non-coordonnée à l'autre lorsque l'on apprend à se connaître. Aussi, les individus peuvent se coordonner pour marcher dans la même direction sans pour autant faire attention à ce que ressent l'autre. De plus, l'importance des trois composantes du rapport varie dans le temps (TICKLE-DEGNEN & ROSENTHAL, 1990) - Figure 3.4. Les comportements caractérisant l'attention mutuelle, la positivité et la coordination devraient toujours être présents, mais à un certain niveau : dès le début lorsque la relation se développe, et plus tard lorsqu'elle est maintenue. Par conséquent, la coordination est moins présente au début de l'interaction, à cause des maladresses et des malentendus dus au manque de connaissance entre les interactants. Cependant, la positivité ne nécessite pas d'expérience préalable avec l'autre interactant et devrait être présente dès le début dans le but de suivre les normes sociales. En plus de la durée de l'interaction, la pondération de chaque composante varie en fonction du contexte de l'interaction, du rôle des participants et de leurs objectifs. Par exemple, l'attention mutuelle peut être l'élément le plus important dans une situation médicale, car il est essentiel d'échanger des informations pour le diagnostic et le traitement.

FIGURE 3.4.

*Importance des trois composantes du rapport à travers le temps (*TICKLE-DEGNEN & ROSENTHAL, *1990).*



La difficulté à créer le rapport varie selon la situation. Dans un contexte académique, les enseignants doivent accorder deux aspects qui semblent contradictoires : être à la fois amical (interpersonnel) et orienté vers la tâche (pédagogique) (NGUYEN, 2007). KYRIACOU (2010) a tout de même défini une classe de cours efficace comme étant certes orientée vers la tâche, mais se caractérisant également par le rapport. Le rapport, dans un contexte académique, implique de faire attention à la progression des étudiants, et en les respectant en tant qu'apprenants et individus. Respecter les élèves en tant qu'apprenants se caractérise par : (1) adapter l'apprentissage aux besoins des élèves, (2) suivre les progrès et apporter son aide de manière patiente et constructive, (3) féliciter, (4) se préoccuper du manque de progrès de l'élève. Respecter les élèves en tant qu'individus se fait indépendamment du contexte académique, et contribue à de bonnes relations.

Nous présentons les trois composantes du rapport dans les sous-sections suivantes de manière générale, ainsi que la façon dont les enseignants et les tuteurs peuvent s'approprier ces composantes dans les classes grâce aux études de GRANITZ et al. (2009).

3.4.1.1. La positivité

La positivité est l'ensemble des comportements qui témoignent de sensations, de ressentis ou d'émotions positives envers les membres de l'interaction et l'interaction elle-même (POTDEVIN, 2020). Elle est reconnue comme une des stratégies puissantes pour maintenir les relations (OGOLSKY & BOWERS, 2013). La positivité peut s'exprimer

au travers des comportements verbaux en étant poli, amusant, joyeux ou encore optimiste ; et des comportements non-verbaux à l'aide de sourires, de hochements de tête, de rires, et de regards portés vers la personne (TICKLE-DEGNEN & ROSENTHAL, 1990).

La positivité se manifeste à travers l'engagement et l'investissement personnel et émotionnel qui découlent de l'interaction. Plus nous éprouvons des émotions positives au cours d'une interaction, plus nous sommes enclins à nous investir personnellement dans celle-ci, cherchant à la prolonger et à la reproduire. BURGOON et LE POIRE (1999) ont montré que des expériences agréables renforcent notre propension à ressentir des émotions positives, influençant ainsi favorablement notre attitude. La perception d'attitudes positives chez l'interlocuteur contribue à notre épanouissement et à notre engagement dans un échange positif. Ainsi, les comportements d'engagement relationnel se traduisent par un intérêt à investir dans l'interaction et une volonté de fournir des efforts pour la maintenir, l'entretenir et la nourrir (LIANG et al., 2009). À long terme, cette notion s'apparente au désir de préserver des relations significatives et précieuses, ainsi qu'à la volonté de faire partie d'un groupe social (HANNA & RICHARDS, 2018).

Dans les classes, les enseignants peuvent adapter la positivité en créant des liens avec les élèves, en se préoccupant d'eux tout en faisant preuve d'empathie. En évitant le favoritisme pour comprendre les besoins des élèves, chaque élève devrait se sentir à l'aise. Cela comprend donc la création d'un environnement positif, convivial et coopératif (GRANITZ et al., 2009).

3.4.1.2. L'attention mutuelle

L'attention mutuelle englobe tous les comportements qui témoignent de la reconnaissance d'autrui, répondant au besoin d'être compris et entendu en tant qu'individu. La nécessité d'écouter et d'être écouté, ainsi que de voir et d'être vu, est très importante dans toute relation et interaction. Le fait de ne pas être écouté et d'être ignoré peut susciter un sentiment de non-appartenance, comme c'est le cas chez les femmes en sciences. L'attention mutuelle concerne d'abord la reconnaissance de l'autre, impliquant la considération de tous les aspects qui définissent une personne afin de mieux comprendre ses ressentis, ses perceptions et ses comportements. Cela signifie également prendre en compte l'individualité de l'autre, le considérant comme une personne à part entière et unique. La notion de reconnaissance de l'autre s'étend également à la reconnaissance des interactions passées et de toutes les informations partagées entre les interactants au cours de celles-ci.

De plus, l'attention mutuelle reflète les sentiments de confiance. Si l'on perçoit l'autre interactant comme attentif et prêt à répondre à nos besoins, cela crée un lien émotionnel (MACINTOSH, 2009). L'expression de comportements rassurants, tels que des stratégies verbales d'empathie, des postures orientées vers l'avant, des expressions faciales positives, des hochements de tête ou encore des regards envers l'autre

personne montre la sensibilité d'un individu à répondre aux besoins de l'autre et renforce le sentiment de sécurité (POTDEVIN, 2020). Enfin, encourager et soutenir l'autre occupe une place particulière dans le rapport, notamment dans le domaine de l'éducation. Cela comprend la disponibilité pour répondre aux questions, fournir des conseils, des commentaires, du soutien, de l'aide et des félicitations lorsque l'élève réussit. L'attention mutuelle arrive à concilier la dimension sociale et professionnelle de l'enseignement. Il faut également être disponible et réactif. De plus, il faut pouvoir mutuellement parler de soi pour prendre en considération les élèves en tant qu'individus, tout en reconnaissant les différences entre les étudiants et les enseignants. Enfin, cela implique également la création d'un climat de confiance par des comportements cohérents et prévisibles, du respect, de la courtoisie, de l'écoute, et en faisant preuve de patience avec les élèves, sans les rabaisser (GRANITZ et al., 2009).

Une autre dimension de l'attention mutuelle est d'être transparent et honnête, en montrant un côté humain aux élèves, et en admettant ses fautes et ses erreurs (GRANITZ et al., 2009). Cela se manifeste à travers la révélation de ses propres pensées ou d'informations personnelles. En outre, il s'agit d'exprimer des sentiments et des ressentis, des comportements de partage d'une évaluation émotionnelle ou d'un jugement d'une situation donnée, ainsi que des comportements d'expression d'une opinion personnelle. Des comportements non-verbaux sont associés lorsque l'on révèle des ressentis ou des informations sur soi-même, tel l'expression faciale d'une émotion ou l'expression d'une gestuelle orientée vers soi (POTDEVIN, 2020).

3.4.1.3. La coordination

Quand les individus deviennent proches, non seulement leur style conversationnel change (e.g., lexique, prosodie) pour s'adapter à l'autre, leur comportement non-verbal change également au cours de l'interaction (CASSELL & TVERSKY, 2005). Cette coordination permet une interaction plus fluide et moins gênante pour les interactants, où les prises de parole sont respectées, ce qui permet une meilleure compréhension. D'après PARK et BURGESS (1924), le rapport implique l'existence d'une réaction coordonnée tel que les interactants réagissent spontanément et de manière sympathique aux attitudes de l'autre. La coordination peut être atteinte de manière verbale en adaptant son lexique ou sa prosodie à son interlocuteur. Elle peut également l'être de manière non-verbale selon deux façons : (1) avec des mouvements synchronisés qui correspondent à un timing précis pour coïncider avec les mouvements de l'autre ; (2) avec des mouvements miroirs, que l'on appelle mimicrie, qui consistent à copier les mouvements de l'autre (BERNIERI, 1988).

Dans le domaine de l'éducation, les enseignants peuvent se coordonner aux élèves en partageant des valeurs, des attitudes, des styles sociaux, des croyances, et des mimétismes comportementaux. Des comportements verbaux ou non-verbaux coordonnés permettent d'établir un lien de similitude entre les interactants. Les individus, adultes comme enfants, ont en effet tendance à se rapprocher de ceux qui leur ressemble

(SHRUM et al., 1988). Comme nous l'avons déjà montré, d'après la théorie de l'apprentissage cognitif social de BANDURA (1986), les individus apprennent souvent en imitant des personnes qu'ils perçoivent comme similaires (ou supérieures en rang ou en statut). Cela implique pour les enseignants de s'engager dans la révélation de soi, et d'adapter son langage corporel avec des mouvements similaires à l'autre (GRANITZ et al., 2009).

La coordination a été montré comme un élément extrêmement important dans le domaine de l'éducation (SINHA, 2016). En analysant un corpus de dialogues entre deux élèves (un tuteur et un apprenant), SINHA et CASSELL (2015) ont montré que la convergence des comportements verbaux ainsi qu'un haut niveau de rapport corrèlent avec une amélioration de l'apprentissage. Les autrices ont expliqué que l'apprentissage étant un processus à la fois cognitif (DILLENBOURG & JERMANN, 2007) et social (KREIJNS et al., 2003), les efforts entre deux élèves pour partager leurs connaissances améliorent l'apprentissage (SCHWARTZ, 1995). Des études ont montré que la coordination entre les interactants est importante pour atteindre ce partage de connaissances (PICKERING & GARROD, 2004).

Le rapport apparaît comme un des éléments essentiels pour la réussite des élèves. En informatique, notamment dans le domaine des agents virtuels, le rapport virtuel a été modélisé et a été étudié à de nombreuse reprise, dans divers domaines d'applications (e.g., recommandation de films, PECUNE et al., 2019, éducation, PAPANGELIS et al., 2014). Dans la prochaine sous-section, nous présentons des études modélisant le rapport virtuel dans les interactions humain-agent ainsi que les défis de la recherche concernant l'intégration du rapport dans ces interactions.

3.4.2. Modélisation du rapport virtuel

Le rapport *virtuel* a été étudié dans les interactions humain-agent dans plusieurs domaines. Il peut être modélisé par des systèmes à base de règles provenant de la littérature en sciences sociales, comme celles présentées dans la sous-section précédente. L'agent virtuel de GRATCH et al. (2006), premier agent simulant le rapport virtuel, écoutait un participant raconter une histoire et adaptait son comportement non-verbal en suivant des règles spécifiques selon le comportement du participant. Il était alors principalement doté de la composante de la coordination. Ce type d'agent, défini par des règles, est caractérisé par un ensemble : (1) d'**actions** par des événements comportementaux; (2) d'**états** en suivant les conséquences des actions (e.g., après avoir détecté "penché à gauche", l'utilisateur est dans l'état "penché à gauche"). Ces états peuvent être contraints sur des bases logiques (e.g., l'utilisateur ne peut pas parler et être silencieux en même temps) ou temporairement (e.g., l'agent reste dans le même état pendant un certain temps) ; (3) d'**animations** par des comportements physiques et pouvant être associé aux actions des agents (e.g., un comportement d'écoute peut être associé à un hochement de tête) ; (4) de **réactions** où les actions de l'utilisateur sont reliés à une ou plusieurs actions de l'agent (e.g., si l'utilisateur hoche

la tête, l'agent hoche la tête). En utilisant ce même agent, N. WANG et GRATCH (2009) ont pu analyser les expressions faciales des utilisateurs pouvant être des indicateurs -ou non- du rapport, et ainsi améliorer l'adaptation de l'agent virtuel. Parmi les signaux non-verbaux, les expressions faciales sont le facteur le plus important dans le jugement du comportement (AMBADY & ROSENTHAL, 1992). Les résultats ont montré que les expressions faciales négatives tel que le dégoût sont des indicateurs d'un manque de rapport, alors que les expressions faciales positives sont au contraire des bons indicateurs. La présence des hochements de tête de l'agent a également amélioré le rapport. De plus, cet agent a été jugé comme étant plus engageant que des humains ayant le même rôle d'écoute. HUANG et al. (2011) ont amélioré le rapport ressenti lors de l'interaction avec l'agent de GRATCH et al. (2006), en augmentant la coordination. En entraînant un modèle de champs aléatoires conditionnels (*conditional random fields* ou CRF) sur 45 vidéos, ils ont pu prédire le moment où l'agent doit hocher la tête et sourire pour un meilleur comportement d'écoute. Les limites de ces études sont que seuls les hochements de tête et les sourires ont été modélisés, et que le contenu du dialogue de l'utilisateur n'a pas été utilisé. Cela implique une absence de coordination au niveau verbal. De plus, les études présentées jusqu'à présent modélisent le rapport pour un comportement d'écoute, en se focalisant plutôt sur la coordination, mais qu'en est-t-il lorsque l'agent est en position de locuteur ?

Des agents virtuels ayant comme principal rôle de produire un discours auront moins de possibilité de s'engager dans une conversation avec l'utilisateur, et ainsi de se coordonner à lui. Il peut alors être difficile de créer une interaction avec un rapport virtuel. Or, dans la théorie de la communication, HORTON et RICHARD WOHL (1956) ont introduit le terme d'*interaction parasociale* pour référer au fait que les individus interagissent souvent avec des médias (télévision, films) comme s'ils étaient engagés dans une vraie interaction sociale. Un agent virtuel ayant principalement un rôle de locuteur pourrait donc créer une relation de rapport virtuel, comme montré par HUANG (2013) selon le principe de l'interaction parasociale. Par exemple, BAILENSON et YEE (2005) ont utilisé des agents délivrant un message persuasif prônant une politique de sécurité sur le campus universitaire. Les auteurs ont utilisé une seule composante du rapport, la coordination à travers le mimétisme, en copiant les hochements de tête des utilisateurs. Ces agents étaient jugés comme étant plus sympathiques et plus persuasifs que les agents avec des comportements scriptés. Dans un contexte médical pour les patients atteints de schizophrénie, RAFFARD et al. (2018) ont également utilisé un agent virtuel copiant les mouvements de tête et du corps des utilisateurs, résultant à un meilleur sentiment de rapport virtuel qu'avec un agent scripté. GORDON et al. (2016) ont quant à eux adapté le comportement de leur robot dans un contexte pédagogique en fonction des expressions faciales des enfants. Ce robot adaptatif a été jugé comme plus engageant et sympathique qu'un même robot dont le comportement a été scripté.

On remarque que, souvent, les études se focalisent sur une seule composante du rapport (e.g., l'attention mutuelle ou la coordination) pour créer des agents capables

de simuler le rapport virtuel. Certaines études ont plus particulièrement étudié le comportement non-verbal des agents virtuels concernant la composante de la positivité. Ainsi, les différents types de sourires (OCHS et al., 2010), les expressions faciales (NUMATA et al., 2020), les attitudes sociales (RISTORCELLI et al., 2023), les postures (CHANG et al., 2019) ou encore les gestes (I. WANG & RUIZ, 2021) ont été étudié pour que les utilisateurs puissent percevoir les interactions avec un agent virtuel comme plus positives. À l'aide de ces études mais également d'analyses de corpus d'interaction humain-humain, il peut alors être possible de prédire automatiquement un niveau de rapport perçu de l'interaction selon le comportement non-verbal de l'utilisateur (MADAIO et al., 2017 ; MÜLLER et al., 2018).

En outre, beaucoup d'études se focalisent sur les comportements non-verbaux des agents virtuels, alors que d'après TICKLE-DEGNEN et ROSENTHAL (1990), le comportement verbal est tout aussi important pour créer le rapport. R. ZHAO et al. (2014) ont ainsi créé un modèle computationnel prenant en compte les stratégies de communication, les objectifs, et la durée d'interaction entre l'agent et l'utilisateur. En analysant un corpus de dialogues entre des élèves lors d'une tâche de tutorat (Figure 3.7 pour des exemples de ce corpus), ils ont décelé des stratégies de communication favorisant le rapport. C'est une de leurs principales contributions : expliquer comment certaines stratégies de communication sont connectées à un but social lié au rapport, et les utiliser lors d'interaction humain-agent (Figure 3.5). Par exemple, partager des informations personnelles ou encore briser les normes sociales après un certain temps permet de croître le sentiment d'attention mutuelle et de coordination. Ce modèle prend en compte la relation entre l'agent et l'utilisateur (e.g., selon la durée d'interaction) et met en évidence comment ces stratégies de communication évoluent selon s'il s'agit d'une interaction entre deux amis ou entre deux étrangers. En analysant le corpus de dialogues, les auteurs ont également montré les stratégies et comportements qui peuvent détériorer le rapport (Figure 3.6). Ceci permet aux chercheurs de connaître les comportements à éviter lors de la construction de modèle de rapport, mais peut également être utile pour des algorithmes de prédiction du niveau de rapport. Bien que R. ZHAO et al. (2014) ont plus mis en avant les comportements verbaux que non-verbaux pour simuler le rapport virtuel, leur modèle permet de couvrir toutes les composantes du rapport (e.g., positivité, attention mutuelle, coordination).

FIGURE 3.5.

Stratégies de communication et fonctions sociales pour construire et maintenir le rapport (R. ZHAO et al., 2014).

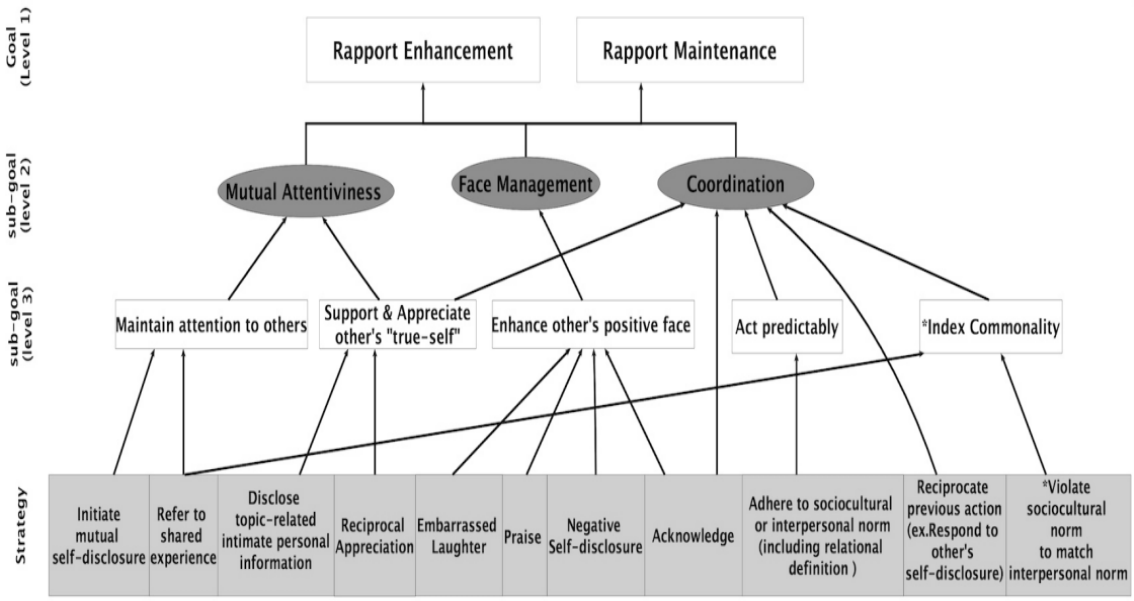


FIGURE 3.6.

Stratégies de communication et fonctions sociales baissant le niveau de rapport (R. ZHAO et al., 2014).

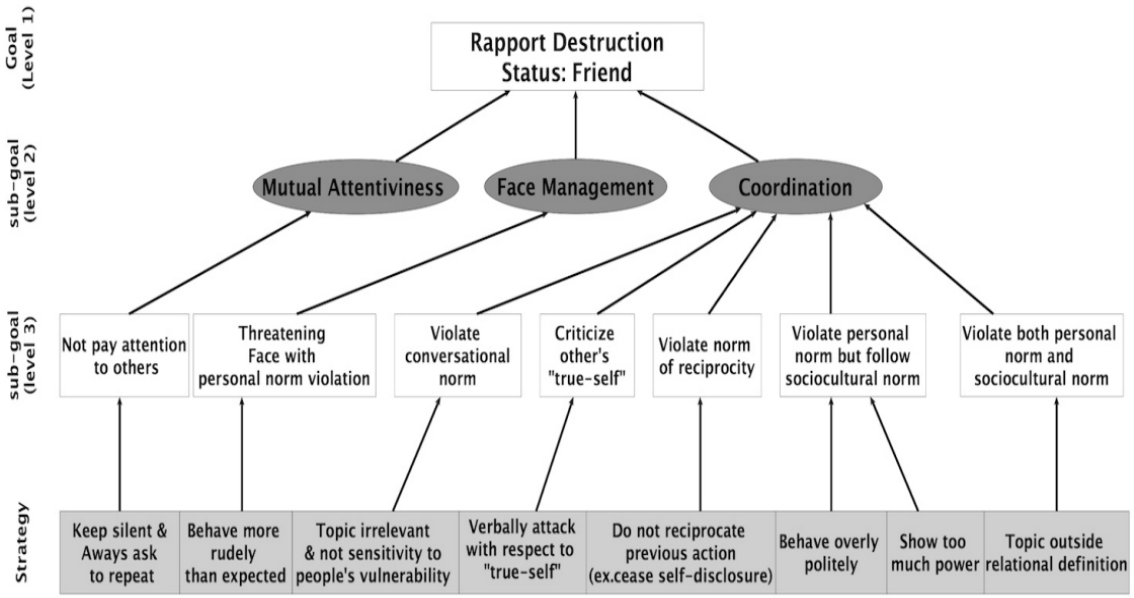


FIGURE 3.7.

Exemples de dialogues entre 2 élèves ne se connaissant pas (à gauche) et entre 2 élèves amis (à droite) provenant du corpus de R. ZHAO *et al. (2014).*

Stranger-Example 1
P1: b equals nineteen over nine
[s_1 = N/A,t_d = 1]
P2: {laughter} good job
[s_2 = praise, t_d = 1]
R=Increase

Stranger-Example 2
P1:I suck at negative numbers
[s_1 = negative self-disclosure, t_d = 1]
P2: it's okay so do I
[s_2 = reciprocate self-disclosure, t_d = 1]
R=Increase

Stranger-Example 3
P1: x equals sixty-four over three
[s_1 = N/A,t_d = 1]
P2: yep
[s_2 = acknowledge, t_d = 1]
R=Increase
P1: x all right thanks .. all right
[s_1 = adhere to sociocultural norm, t_d = 1]
P2: it was a complicated one
[s_2 = face-boosting acknowledgment, t_d = 1]
R=Increase

Friend-Example 1
P1: are there any girls you like
[s_1 = elicit self-disclosure, t_d = 3
more personal topic]
P2: all of them are not the best looking
[s_2 = reciprocate self-disclosure,t_d = 3]
R=Increase

Friend-Example 2
P1:remember you went to Connecticut
[s_1 = Refer to shared experience,t_d = 2]
P2:that was just to visit my cousin
[s_2 = disclose topic-related intimate
personal information, t_d = 2]
R=Increase

Friend-Example 3
P1: silly goose that's a backwards two
[s_1 = violate sociocultural norm to adhere
to interpersonal norm, t_d = 1]
P2:two
[s_2 = N/A, t_d = 1]
R=Increase

Note. Un élève jouait le rôle du tuteur pendant que l'autre jouait le rôle de l'apprenant. Ce corpus a permis aux auteurs d'analyser les stratégies s favorisant ou baissant le rapport R. t_d est la profondeur de l'arbre de la Figure 3.5.

Z. ZHAO et al. (2018) ont alors étudié le rôle du rapport et des stratégies de communication dans l'apprentissage. De la même manière que l'étude de SINHA et CASSELL (2015) et R. ZHAO et al. (2014), les auteurs ont analysé un corpus de dialogues entre deux élèves, l'un jouant le rôle de tuteur, l'autre d'apprenant. Chaque discours a été labelisé avec une stratégie d'enseignement (e.g., proposer une réponse, demander de l'aide) et une stratégie de communication (e.g., féliciter, révéler des informations personnelles). Les dialogues ont été labelisés avec un niveau de rapport (faible, moyen, haut) et un gain d'apprentissage (e.g., problème correctement résolu). Les auteurs ont comparé deux modèles pour prédire les meilleures stratégies de communication et d'enseignement qu'un agent virtuel devrait employer pour améliorer l'apprentissage de l'élève. En comparant un LSTM et un modèle de processus markovien, les auteurs ont montré que le LSTM prenant en compte le niveau de rapport, ainsi que l'historique des stratégies de communication et d'enseignement utilisées entre un tuteur et un apprenant, permettait de prédire de meilleures décisions pour améliorer l'apprentissage. Cette étude montre que les facteurs sociaux ont un fort effet positif sur la performance d'un apprenant dans une tâche pédagogique. Ce modèle n'a cependant pas été déployé et évalué dans un agent virtuel. De plus, les comportements

non-verbaux ne sont pas pris en compte dans ce modèle.

Une des limites que l'on rencontre dans la plupart des modèles computationnels de rapport virtuel est que les agents virtuels donnent l'illusion de comprendre ce que l'utilisateur dit, sans aucun module de traitement automatique du langage. Dans un contexte médical avec des patients atteints de dépression, DEVAULT et al. (2014) ont créé un agent virtuel avec des comportements de rapport virtuel, pouvant détecter des signaux non-verbaux du stress des patients, avec un système de dialogue existant dans la littérature (MORBINI et al., 2014) pour converser avec les patients. Grâce à cela, l'agent a pu poser des questions en lien avec ce que les patients racontaient. De plus, l'agent virtuel a pu donner plus de retour lors du *back-channeling* (e.g., l'écoute active), en répondant par exemple par « je suis désolée » ou « c'est super » selon ce que le système de dialogue a reconnu. Les comportements de rapport virtuel de cet agent, mais également le fait de parler à un agent virtuel, ont permis aux patients de se sentir plus à l'aise, tout en réduisant leur peur d'être jugé.

Si l'on applique directement les résultats des études de TICKLE-DEGNEN et ROSENTHAL (1990) en sciences sociale sur les agents virtuels, trois composantes (e.g., positivité, coordination, et attention mutuelle) devraient augmenter le sentiment de rapport virtuel entre un agent virtuel et un utilisateur. LUGRIN et al. (2022) ont considéré un quatrième élément, spécifique aux interactions avec des agents virtuels : l'*anonymat*, le sentiment que notre identité est préservée. Puisque l'anonymat n'est pas possible lors d'interaction en face-à-face avec des humains, cette composante n'était pas pertinente pour les études de TICKLE-DEGNEN et ROSENTHAL (1990). Manipuler un sentiment d'anonymat peut encourager des effets positifs du rapport, comme révéler plus d'informations sur soi-même. Des recherches montrent en effet qu'un plus grand sentiment de rapport permet aux personnes de plus parler d'elle-même (GRATCH et al., 2007). Les agents virtuels, en tant que programmes informatiques à priori non contrôlés par un humain, permettent aux individus de ressentir un certain anonymat leur permettant de se sentir plus à l'aise et de révéler des informations personnelles, sans crainte d'être jugés (DEVAULT et al., 2014) ou discriminés (VILARO et al., 2021). De plus, d'autres études ont montré que c'est bien la combinaison du rapport virtuel et du sentiment d'anonymat procuré par l'agent virtuel qui permet aux utilisateurs de révéler plus d'informations importantes sur eux, et pas seulement l'anonymat. Les participants des études de LUCAS et al. (2017) ont révélé plus d'informations sur leur santé mentale lorsqu'ils ont parlé avec un agent virtuel avec des comportements de rapport, que lorsqu'ils ont répondu à un questionnaire anonyme portant sur les mêmes questions. Cela pose tout de même des questions éthiques, notamment sur le troisième principe de l'Organisation de Coopération et de Développement Économiques sur la transparence et la responsabilité des systèmes d'intelligence artificielle. Avec un groupe d'étudiants, SCHLIMBACH et KHOSRAWI-RAD (2022) ont enrichi ce principe avec les règles suivantes : ce genre de modèle d'agent virtuel pédagogique doit démontrer comment les données des utilisateurs sont analysées, se conformer aux Règlement Général sur la Protection des Données (RGPD), et ne pas abuser de la

confiance qu'un utilisateur pourrait accorder à l'agent virtuel. Modéliser le rapport virtuel reste à ce jour un défi dans le domaine des agents virtuels, que ce soit du côté éthique ou du côté computationnel (CLAVEL et al., 2022).

3.5. Vers un modèle d'agent virtuel pédagogique pour lutter contre la menace du stéréotype

La revue systématique présentée dans ce chapitre, ainsi que la présentation des travaux sur les agents virtuels pédagogiques s'approchant des modèles de réussite, et des travaux sur l'impact positif du rapport dans l'éducation et des modélisations computationnelles existantes du rapport, nous permettent de tirer des conclusions en répondant à chacune des questions posées dans la Section 3.2.1.2.

Les individus perçoivent-ils différemment les agents virtuels selon leur propre genre et celui de l'agent ? Comme les individus communiquent avec les agents virtuels en appliquant des règles sociales et des attentes en tant qu'êtres sociaux (HEYSELAAR et al., 2023 ; NASS & MOON, 2000), il n'est pas surprenant qu'ils appliquent également des stéréotypes de genre aux agents virtuels et à leurs interactions avec eux (Section 3.2.3). Les agents virtuels féminins sont généralement considérés comme moins experts, moins compétents et moins puissants que les agents virtuels masculins (BAYLOR & KIM, 2004 ; NUNAMAKER et al., 2011), et ils sont également généralement perçus comme plus sympathiques et attrayants que les agents virtuels masculins (LUNARDO, BRESSOLLES et al., 2016 ; NUNAMAKER et al., 2011). Ces différences de perception peuvent même affecter les décisions des individus, où les participants hommes ont tendance à faire plus confiance à un agent virtuel masculin (ŚWIDRAK et al., 2021).

Au vu des résultats empiriques, nous proposons de répondre simultanément à la question 2, **Le genre des agents virtuels pédagogiques influence-t-il les performances ainsi que d'autres facteurs éducatifs (motivation, sentiment de réussite, ...) des élèves ?**, et à la question 3, **Y a-t-il des tâches ou des situations académiques auxquelles un agent virtuel masculin est mieux adapté qu'un agent virtuel féminin, et vice versa, selon des études empiriques ?**. La revue systématique menée dans la section 3.2 pourrait laisser croire que les agents virtuels pédagogiques masculins sont mieux adaptés que les agents virtuels pédagogiques féminins pour améliorer les résultats scolaires, en particulier dans les domaines scientifiques à prédominance masculine comme les domaines des STIMs (MAKAROVA et al., 2019). Cependant, la recherche montre également que les agents virtuels pédagogiques féminins peuvent améliorer les performances des apprenants dans ces domaines, en particulier lorsque ces agents se présentent comme étant compétents dans le domaine (PLANT et al., 2009 ; R. ROSENBERG-KIMA et al., 2008). Certaines études ont ainsi montré que l'utilisation d'agents virtuels féminins comme modèles de réussite dans le domaine de l'ingénierie peut augmenter le sentiment de réussite des participantes (R. ROSENBERG-KIMA et al.,

2008), et l'intérêt des participants, hommes comme femmes (R. B. ROSENBERG-KIMA et al., 2010) et filles comme garçons (PLANT et al., 2009), ainsi que leurs performances (PLANT et al., 2009). Ces résultats sont particulièrement pertinents pour lutter contre les effets de menace du stéréotype, un phénomène qui illustre comment et pourquoi les performances des étudiantes en mathématiques peuvent être altérées par les stéréotypes de genre (Chapitre 2). Des recherches antérieures en psychologie sociale ont montré que les effets de menace du stéréotype peuvent être contrecarrés en introduisant un modèle de réussite représenté par une femme ou une fille qui ressemble aux apprenantes afin qu'elles puissent s'identifier à ce modèle (BAGÈS et al., 2016). Un tel modèle positif pourrait être incarné par un agent virtuel pédagogique (R. ROSENBERG-KIMA et al., 2008), et utilisé pour réduire les effets de menace du stéréotype dans les domaines des STIMs.

Comment le rôle d'un agent virtuel impacte-t-il ces résultats ? Les individus ont tendance à écouter davantage les agents virtuels dont le genre correspond de manière stéréotypée au contexte ou au rôle, tels que les agents virtuels féminins avec les cosmétiques et les agents virtuels masculins avec le sport (E.-J. LEE, 2003), les agents virtuels féminins dans des contextes impliquant une influence sociale (KHASHE et al., 2017), et les agents virtuels féminins dans un rôle d'assistante (ESPOSITO et al., 2021). Le rôle des agents virtuels pédagogiques a également été étudié dans un contexte académique, à l'aide d'agents *expert*, *motivateur* et *mentor* (BAYLOR & KIM, 2004). Les agents experts sont plus âgés que les étudiants, autoritaires, strictement informatifs et cultivés. Les agents motivateurs sont enthousiastes et ne sont pas considérés comme particulièrement cultivés, ils sont principalement utilisés pour susciter la motivation. Quant aux agents mentors, ils sont légèrement plus âgés que les étudiants, sont cultivés et sont également utilisés pour susciter la motivation. Dans l'expérience de BAYLOR et KIM (2005), l'agent motivateur a contribué à accroître le sentiment de réussite des étudiants, tandis que l'agent expert a permis aux étudiants d'avoir de meilleur résultat sur une tâche de planification de cours. L'agent mentor semble être le plus efficace, car il a renforcé le sentiment de réussite des étudiants tout en améliorant leur performance, en plus d'avoir été jugé comme étant le plus engageant. Ces résultats correspondent aux attentes des étudiants et des enseignants dans les interactions humain-humain (BEISHUIZEN et al., 2001) et agent-humain (Y. KIM, 2007) : un tuteur compétent dans son domaine offrant du soutien. L'agent mentor s'apparente à un modèle de réussite pouvant créer le rapport virtuel. Ainsi, un agent virtuel pédagogique féminin conçu comme étant un agent mentor pourrait améliorer la performance des apprenantes (PLANT et al., 2009), mais un agent virtuel pédagogique féminin conçu comme un agent motivateur pourrait ne pas être efficace sur les performances des apprenantes (SHIBAN et al., 2015). L'impact du genre de l'agent virtuel dépend de plusieurs aspects liés au rôle de l'agent. Ce rôle se reflète en effet à travers l'apparence mais aussi le discours et les gestes de l'agent.

Comment l'apparence et les capacités interactives d'un agent virtuel impactent-elles ces résultats ? Les sujets des interactions des individus avec les agents virtuels

diffèrent selon leur apparence genrée : dans les études de BRAHNAM et DE ANGELI (2012) et DE ANGELI et BRAHNAM (2006), les agents virtuels féminins ont reçu significativement plus de propositions sexuelles violentes, plus de menaces de viol, plus de commentaires sur leur apparence, et plus d'insultes par rapport aux agents virtuels masculins qui ont reçu peu de propositions sexuelles, la plupart ciblant leurs petites amies. De plus, le degré de masculinité et de féminité perçu peut influencer la décision des hommes, comme l'a montré ŚWIDRAK et al. (2021) où les participants ont été davantage convaincus par des agents d'aspect masculin que par des agents d'aspect moins masculin, quel que soit le genre des agents. Comme précisé dans la question précédente, le rôle des agents virtuels est important dans les situations académiques et l'apparence des agents virtuels peut influencer leur rôle perçu : un agent virtuel d'apparence plus âgée portant une cravate pourrait être perçu comme un expert, contrairement à un jeune agent virtuel (BAYLOR & KIM, 2005 ; Y. KIM & BAYLOR, 2016 ; SHIBAN et al., 2015). Outre le rôle et l'apparence des agents, la recherche a montré l'importance du rapport virtuel entre les apprenants et les agents pédagogiques (KRÄMER et al., 2016). Différents modèles de rapport virtuel existent à ce jour (Section 3.4.2) et pourraient être utilisés pour réduire la menace du stéréotype. En effet, une des composantes du rapport étant la coordination, cela engendrerait des similitudes entre l'apprenant et l'agent virtuel. Les similitudes sont importantes pour percevoir une autre personne comme un modèle de réussite. Les autres composantes du rapport, la positivité et l'attention mutuelle, permettent d'établir un climat de confiance où les élèves se sentent considérés et compris. Il est important pour les élèves de se sentir écouter. En effet, des recherches en sciences sociales ont montré que les élèves, et en particulier les filles, croient plus en leur capacités mathématiques lorsque les professeurs s'intéressent à eux en tant qu'individus (FORGASZ & LEDER, 1996). Ces recherches tendent à montrer qu'un modèle de réussite féminin, incarné par un agent virtuel pédagogique capable d'établir une relation positive avec les apprenants à travers le rapport virtuel, peut réduire les effets de menace du stéréotype, et ainsi améliorer la performance, l'intérêt et le sentiment de réussite des filles en mathématiques.

Les agents virtuels androgynes sont-ils une solution potentielle pour lutter contre les stéréotypes de genre ? Cette question est assez difficile à répondre car à notre connaissance, peu de recherches ont exploré l'utilisation d'agents virtuels androgynes pour combattre les stéréotypes de genre en contexte académique. SILVERVARG et al. (2013) ont utilisé un agent virtuel pédagogique androgyne dans un contexte éducatif et ont montré que les participants, bien qu'évaluant l'agent comme « pas clairement un garçon ni une fille », ont tendance à attribuer un genre binaire (garçon ou fille) à l'agent. Les auteurs ont donc prudemment supposé que les élèves pouvaient attribuer leur propre genre à un agent virtuel androgyne, leur donnant ainsi plus de liberté et faisant de l'agent un modèle de réussite approprié pour chacun. Cependant, les résultats n'ont pas montré quel genre les participants ont attribué à l'agent, ni pourquoi. La façon dont les individus attribuent un genre à un agent virtuel androgyne ou asexué devrait être davantage étudiée. MCDONNELL et al. (2009) et ZIBREK et al. (2020) ont

par exemple montré que l'attribution d'un genre à des agents virtuels neutres (e.g., des mannequins en bois) dépendait de leurs démarchent stéréotypiquement associés aux femmes ou aux hommes. NICULESCU et al. (2009) ont quant à eux montré que le genre attribué à un agent virtuel androgyne dépendait de l'agent virtuel que les participants ont vu juste avant. Nous pouvons également nous demander si l'attribution de genre dépend du contexte ou du rôle des agents virtuels. Étant donné que les domaines des STIMs sont considérés comme étant des domaines masculins (MAKAROVA et al., 2019), les participants pourraient penser que les agents virtuels androgynes sont des hommes. Cela pourrait renforcer le stéréotype selon lequel les domaines des STIMs sont plus adaptés aux hommes qu'aux femmes. En ce qui concerne les rôles des agents, il y a plus d'assistantes virtuelles femmes que d'assistants virtuels hommes. Certains développeurs ont admis que les assistantes virtuelles femmes sont généralement utilisées car elles évoquent des stéréotypes de genre : les femmes sont censées servir, aider et entretenir les autres (BRAHNAM & WEAVER, 2015) ; bien que d'autres études ont montré que les robots à l'apparence féminine et les assistants vocaux à la voix féminine sont considérés comme plus humain (BORAU et al., 2021). Cependant, les assistants virtuels androgynes pourraient être considérés comme des femmes, et renforcer les stéréotypes néfastes sur les femmes.

3.6. Discussion

Les agents virtuels pédagogiques sont connus pour améliorer la performance des élèves dans un environnement virtuel d'apprentissage. En se fondant sur les recherches présentées dans ce chapitre, non seulement le genre des agents virtuels pédagogiques peut être important, mais également leur comportement verbal et non-verbal, leur rôle, et leur origine culturelle. Les filles pourraient percevoir un agent virtuel pédagogique féminin comme un modèle de réussite qui influence leur motivation et leurs efforts pour apprendre. Des mesures d'identifications à ces agents virtuels n'ont cependant pas été effectué. Néanmoins, les effets positifs sont là : des agents virtuels féminins, se présentant comme ingénieures, ont pu augmenter la motivation et l'intérêt en sciences des participantes. Les agents virtuels pourraient être utilisés pour incarner des modèles de réussite et ainsi modifier les attitudes et la motivation de l'apprenant. Comme décrit au chapitre 2, un modèle de réussite incarné par une femme ou une fille qui réussit en mathématiques peut réduire les effets de menace du stéréotype. On peut alors se demander quel serait l'impact d'utiliser systématiquement un agent virtuel du même genre que les apprenants ?

N'utiliser que des modèles de réussite virtuels masculins performants en mathématiques avec des garçons pourrait renforcer les stéréotypes de genre. Les femmes sont conscientes du stéréotype négatif sur leurs compétences en mathématiques qui crée un environnement hostile pour elles. Une solution, au moins dans les domaines des STIMs, serait d'utiliser des modèles de réussite virtuels féminins expliquant comment bien réussir en mathématiques (Section 2.3.4.3). Puisque le manque de modèles de

réussite incarnés par des femmes peut amener les étudiantes à croire que les hommes sont meilleurs que les femmes dans les domaines des STIMs, il pourrait être intéressant de combiner les résultats des différentes recherches menées dans les domaines de la psychologie cognitive et sociale et des agents virtuels. Présenter des agents virtuels pédagogiques féminins, incarnant un modèle de réussite en mathématiques, pourrait habituer les élèves à voir plus de filles ayant d'une part des bonnes notes en mathématiques, et d'autre part une confiance en ses compétences mathématiques. Ceci pourrait altérer les associations automatiques stéréotypés en mémoire. Les individus, adultes comme enfants, tendent en effet à associer implicitement les hommes avec les sciences et les femmes avec les lettres. Avec plus de modèles de réussite féminins, les élèves pourraient associer à la fois les hommes et les femmes avec les sciences. De plus, notre hypothèse est que la menace du stéréotype peut être réduite lorsqu'un agent virtuel pédagogique incarne un modèle de réussite pour l'apprenant avec une similarité en termes d'apparence et de genre, et dans la condition où cet agent virtuel démontre un comportement de rapport. À ce jour, aucune étude n'a testé une telle hypothèse. Plusieurs travaux de recherche portent sur le développement et l'impact des agents virtuels pédagogiques à des fins éducatives pour améliorer l'apprentissage et les performances des apprenants. Cependant, aucun agent virtuel pédagogique n'a été développé pour lutter contre la menace du stéréotype, à l'origine notamment d'une inégalité de performance genrée en mathématiques (Chapitre 2). Dans le prochain chapitre, nous présentons comment nous pouvons articuler ces deux littératures pour créer et évaluer de tels agents virtuels pédagogiques.

3.7. Points clés du chapitre

- Les agents virtuels pédagogiques peuvent améliorer l'apprentissage selon le rôle qu'ils incarnent (e.g., encourager et montrer ses connaissances).
- Persistance des stéréotypes de genre dans les interactions humains-agents (e.g., les agents féminins sont perçus comme plus gentils, les agents masculins sont perçus comme plus intelligents).
- Pas de résultats tranchés sur l'impact du genre de l'agent sur les performances académiques.
- Très peu d'études ont montré que les agents virtuels pédagogiques créés dans l'objectif de simuler un modèle de réussite pouvaient encourager les femmes en mathématiques.
- Le rapport virtuel (e.g., comportements verbaux et non-verbaux pour créer une interaction positive entre humain-agent) peut améliorer les performances des apprenants.
- Modèles computationnels existants sur le rapport virtuel : copie des mouvements de l'utilisateur, comportements d'écoute, stratégies de conversations favorisant le rapport virtuel, système de dialogue pour poser des questions en lien avec le discours de l'utilisateur.
- Modéliser le rapport virtuel reste à ce jour un défi.
- Notre objectif est de créer un modèle computationnel de rapport virtuel, ainsi que des agents virtuels pouvant incarner des modèles de réussite en mathématiques.

4. Problématique de recherche

L'état de l'art sur les effets de menace du stéréotype a permis de comprendre comment et pourquoi les stéréotypes de genre peuvent conduire les filles et les femmes à obtenir des performances suboptimales dans les domaines des STIMs. Ces effets sont observables lors de tests standardisés difficiles, aussi bien chez les adultes que chez les enfants, et sont encore plus prégnants chez les filles et les femmes qui s'identifient fortement aux mathématiques (C. M. STEELE, 1997) ou qui manifestent un fort ancrage du stéréotype en mémoire sémantique (GALDI et al., 2014). Cette revue de la littérature a également montré que l'enjeu de la recherche a été de mettre en place des interventions susceptibles de réduire ces effets de menace du stéréotype, et d'en identifier les conditions d'efficacité. Parmi les différentes interventions testées et potentiellement applicables en salle de classe, l'exposition à un modèle féminin de réussite en mathématiques s'est avérée efficace dans plusieurs études (LEROY et al., 2022 ; MCINTYRE et al., 2011) dès lors que les filles/femmes peuvent s'identifier à ce modèle (MCINTYRE et al., 2003). D'autres études ont montré que les modèles de réussite masculins pouvaient également améliorer les résultats des filles et des femmes, mais sous certaines conditions comme l'explication de la réussite par le travail et les efforts (BAGÈS & MARTINOT, 2011). En revanche, si le modèle de réussite est perçu comme inaccessible (BAGÈS et al., 2016), menaçant (MARX et al., 2013), ou anxieux vis-à-vis de ses aptitudes en mathématiques (BEILOCK et al., 2010), alors l'effet de menace du stéréotype persiste. Par conséquent, le comportement verbal et non-verbal du modèle de réussite, ainsi que sa capacité à générer de l'identification, doivent faire l'objet d'une attention particulière. Le discours des modèles intervenant dans les classes pourrait en effet décourager les filles dès lors que leur parcours scolaire semble inaccessible, par exemple (BREDA et al., 2021).

Cependant, il est impossible de contrôler toutes ces caractéristiques avec de vrais modèles de réussite. Il n'est pas possible non plus de faire intervenir de vrais modèles dans toutes les classes, et encore moins juste avant de passer un test de mathématiques. Or, l'interaction entre les modèles de réussite et les élèves est importante pour favoriser l'identification (BUCK et al., 2008). On remarque cependant que cette interaction est absente de la procédure qui consiste à faire lire un texte aux élèves sur l'histoire personnelle du modèle, que l'on retrouve dans la majorité des études (BAGÈS et al., 2016). L'enjeu est donc de pouvoir proposer des modèles de réussite interactifs, dont il est possible de contrôler à la fois le comportement verbal et non-verbal pour répondre à toutes les conditions susceptibles de favoriser une identification avec les élèves, et utilisables aux moments opportuns, dans toutes les classes. C'est l'objectif de cette thèse.

Les personnages virtuels utilisés dans les environnements numériques d'apprentissage pourraient incarner des modèles interactifs de réussite efficaces pour réduire les effets de menace du stéréotype. La revue de la littérature en informatique sur les agents virtuels pédagogiques (Chapitre 2) a montré qu'ils pouvaient améliorer la motivation (R. ROSENBERG-KIMA et al., 2008 ; R. B. ROSENBERG-KIMA et al., 2010) et la performance des apprenants (PLANT et al., 2009), et qu'il est possible de créer une forme d'interaction positive entre l'agent et l'apprenant appelé rapport virtuel (GRATCH et al., 2006) qui conduit à de meilleures performances en mathématiques (KRÄMER et al., 2016). La capacité de l'agent virtuel à instaurer un rapport virtuel dépend de trois composantes durant ses interactions avec l'apprenant (HUANG et al., 2011 ; TICKLE-DEGNEN & ROSENTHAL, 1990) : 1) attitudes positives (e.g., acquiescement, sourires), 2) attention mutuelle (e.g., regard mutuel), et 3) coordination (e.g., synchronie, mimétisme). On voit bien ici tous les avantages potentiels liés à l'utilisation d'agents virtuels et à leur déclinaison possible dans un but de réduire les effets de menace du stéréotype : il est possible de contrôler leur apparence, leur discours, leurs gestes ainsi que le moment de leur utilisation au sein de la classe (durant l'apprentissage ou la révision d'une leçon, avant un test, etc). Toutefois, les études réalisées dans le domaine des agents virtuels pédagogiques ont porté en très grande majorité sur des apprenants adultes et des agents virtuels à l'apparence d'adulte. Ceci soulève donc la question de la possibilité de créer des agents virtuels interactifs, adaptés à l'âge et au genre des élèves, susceptibles de créer un rapport virtuel avec les élèves et de favoriser l'identification. De plus, il n'existe pas d'étude, à ce jour et à notre connaissance, portant sur des agents virtuels pédagogiques qui posséderaient les caractéristiques requises pour servir de modèles de réussite en mathématiques pour réduire les effets de menace du stéréotype chez les filles, à savoir : (1) expliquer que sa réussite en mathématiques provient de ses efforts et son travail, (2) avoir une réussite en mathématiques et un statut accessibles, (3) être non-menaçant, (4) ne pas être anxieux quant à ses aptitudes en mathématiques, (5) ne pas véhiculer l'idée de compétences innées en mathématiques.

Les environnements virtuels étant de plus en plus utilisés en classe comme à la maison, il est important de s'intéresser aux personnages virtuels présents dans ces environnements pour faciliter l'apprentissage des mathématiques notamment en proposant un contexte d'apprentissage qui permet de réduire les effets de menace du stéréotype. En liant ces deux ensembles de travaux, d'une part en psychologie sociale sur la menace du stéréotype et d'autre part en informatique sur les agents virtuels pédagogiques, cette thèse a pour objectif d'étudier dans quelle mesure un agent virtuel pédagogique peut réduire les effets de menace du stéréotype chez les élèves de 6e et de 5e en mathématiques en situation de classe habituelle. Nous faisons l'hypothèse que ces agents pourraient réduire les effets de menace du stéréotype chez les filles, et améliorer ainsi leurs performances en mathématiques, et ainsi aider les enseignants face à l'hétérogénéité de leur classe. Nous avons choisi des élèves de 6e et de 5e pour deux raisons. Tout d'abord, parce que l'effet de menace du stéréotype chez les filles en mathématiques est bien documenté au niveau collège dans plusieurs

pays, y compris en France (HUGUET & REGNER, 2007 ; RÉGNER et al., 2014). Ensuite, parce que l'utilisation de l'informatique dans l'apprentissage des mathématiques commence à se généraliser au collège, ce qui justifie d'autant plus de faire attention à ce que ces outils ne reproduisent pas les stéréotypes et/ou leurs effets.

Le premier objectif de cette thèse est de développer des agents virtuels pédagogiques en 3D, adaptés à des élèves de 6e et 5e en fonction de leur genre, et capables d'interactions verbales et non-verbales pour créer un "rapport virtuel" avec les élèves, tout en expliquant leur réussite en mathématiques par leur travail et leurs efforts. Pour permettre l'identification des élèves, les agents filles et garçons doivent avoir l'apparence d'adolescents et ne présenter aucun stéréotype lié aux STIMs. De tels agents englobant toutes les caractéristiques décrites ci-dessus n'existent pas à ce jour et à notre connaissance. Il faut donc les modéliser, les développer, et les évaluer auprès de collégiens. Nous avons analysé les perceptions des élèves envers nos agents selon les principales caractéristiques clés pour établir la possibilité d'une identification entre l'élève et l'agent. Ces caractéristiques comprennent leur apparence, le sentiment de rapport virtuel perçu lors de l'interaction, leur niveau d'anthropomorphisme, leur sympathie, leur niveau de compétence en mathématiques, ainsi que le niveau d'identification à ces agents. Ceci devrait s'accompagner d'une plus grande perception de la sympathie, et d'une plus grande identification. Nous devons également nous assurer que les élèves n'appliquent pas de stéréotype de genre aux agents, et que les agents féminins et masculins sont perçus comme tout aussi compétent en mathématiques l'un que l'autre. Toutes ces évaluations permettront de vérifier si nos agents possèdent les prérequis nécessaires pour réduire les effets de menace du stéréotype.

Le second objectif est de tester si ces agents virtuels peuvent réduire l'effet de menace du stéréotype sur les performances des collégiennes en mathématiques, en comparant l'impact de plusieurs agents virtuels pédagogiques sur la performance des élèves lors d'un test de mathématiques standardisé. Chaque agent rappellera les notions clés d'une leçon de mathématiques, mais se différencie par leur genre (féminin/masculin) et leurs comportements verbaux et non-verbaux. Cette étude ne consiste ni à comparer l'efficacité de nos agents virtuels à de vrais modèles de réussite, ni à les comparer aux histoires personnelles des modèles comme par exemple utilisées par BAGÈS et al. (2016). En vue de l'utilisation accrue du numérique éducatif, il s'agit ici de se concentrer sur les environnements virtuels d'apprentissage, et savoir s'ils sont tous bénéfiques sur la performance. En d'autres termes, le second objectif de cette thèse est d'évaluer les caractéristiques des agents virtuels pédagogiques permettant de réduire les effets de menace du stéréotype chez les filles en mathématiques. Le meilleur modèle de réussite en vue de cet objectif serait un agent virtuel qui explique sa réussite par le travail de manière interactive en induisant un rapport virtuel avec les élèves (comparativement à un agent qui n'explique pas sa réussite ou qui l'explique certes par le travail mais sans avoir créé de rapport virtuel avec l'élève). Ces effets bénéfiques d'un tel agent sur les filles ne devraient pas s'accompagner par une baisse, en parallèle, des performances en mathématiques des garçons. Comme montré par

BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016), les garçons peuvent également bénéficier des modèles de réussite dès lors qu'ils expliquent leur réussite en mathématiques. Nous examinerons également l'effet des agents virtuels pédagogiques sur le sentiment de réussite en mathématiques et l'importance accordée dans ce domaine. Il sera également examiné, de manière exploratoire, si les effets des agents virtuels dépendent de la force d'ancrage des stéréotypes de genre en mémoire sémantique.

La prochaine partie de cette thèse est consacrée à l'atteinte de ces deux objectifs. Après un travail informatique de création, de modélisation, et de développement des agents virtuels pédagogiques ainsi que d'un environnement virtuel d'apprentissage en mathématiques, deux études pilotes réalisées auprès de 144 élèves de 6e et 5e permettront de vérifier que les élèves perçoivent les agents virtuels comme des modèles de réussite en mathématiques, s'identifient à eux, et distinguent les agents selon qu'ils instaurent ou non un rapport virtuel. Une troisième étude pilote auprès de 38 élèves de 6e et 5e permettra de jauger la difficulté des exercices proposés lors du test standardisé de mathématiques de l'étude principale. Cette étude principale permettra ensuite d'évaluer l'impact de nos agents virtuels pédagogiques sur les performances en mathématiques des filles. Au total, 326 élèves de 6e et 5e, guidés par un agent virtuel pédagogique dans leur révision de notions mathématiques, vont passer un test sur ces notions, dans un plan expérimental en 2 (genre de l'agent : féminin/masculin) x 2 (sexe de l'élève : fille/garçon) x 3 (type de l'agent : modèle sans explications et sans rapport/modèle avec explications et sans rapport/modèle avec explications et avec rapport). Toutes les études ont été réalisées directement en classe, en situation quasi-habituelle. Pour éviter les biais méthodologiques tels que l'absence de pression évaluative lors du test, les élèves ont été avertis que les tests de mathématiques compteront dans leur moyenne. Les élèves dont les parents ont refusé la participation à l'étude ont passé les tests de mathématiques sur feuille, en même temps que les autres. Il a de plus été précisé que ces tests étaient des bilans régionaux mandatés par les chefs d'établissement eux-mêmes.

Pour conclure, cette thèse met en lien les recherches provenant de deux domaines, l'informatique et la psychologie sociale, afin d'étudier comment un agent virtuel pédagogique peut réduire les effets de menace du stéréotype chez les filles en mathématiques. Au niveau théorique, cette thèse est susceptible d'apporter une compréhension plus approfondie des utilisations des modèles de réussite pour réduire les effets de menace du stéréotype. De plus, cette thèse aspire à apporter une lumière inédite sur les interactions entre les agents virtuels pédagogiques et les collégiens en vue de réduire les effets de menace du stéréotype, soulignant ainsi l'utilisation de la technologie dans l'éducation. Au niveau applicatif, il sera envisagé de réaliser un prototype d'application de notre environnement virtuel d'apprentissage afin de paramétrer les agents selon d'autres leçons et exercices de mathématiques. Les résultats obtenus de l'étude principale devraient proposer des stratégies et des outils d'apprentissage numérique, contribuant ainsi à l'amélioration de l'expérience éducative des apprenants. De plus, en considérant la place grandissante du numérique éducatif à

la maison comme en classe, cette thèse propose des pistes novatrices pour intégrer efficacement les agents virtuels dans le processus d'enseignement, offrant un soutien individualisé aux élèves. Enfin, cette thèse souligne l'importance de la collaboration entre chercheurs et éducateurs pour fournir aux enseignants des outils numériques pour mieux comprendre et soutenir leurs élèves, tandis que les parents pourraient bénéficier d'une compréhension approfondie des défis auxquels leurs enfants, et plus particulièrement leurs filles, peuvent être confrontés dans le domaine des mathématiques. Cette thèse s'inscrit pleinement dans les objectifs du programme Ampiric, à savoir améliorer la compréhension des processus d'enseignement-apprentissage des savoirs fondamentaux à travers des recherches fondamentales, et coproduire des outils et des applications au service des apprentissages fondamentaux.

Deuxième partie

Études empiriques

5. Introduction de la deuxième partie

La deuxième partie de cette thèse est consacrée aux contributions empiriques. Notre objectif était de développer un agent virtuel pédagogique pour lutter contre les effets de la menace du stéréotype en mathématiques. Pour se faire, nous avons pris en compte les revues sur la menace du stéréotype et sur la modélisation du genre, du rôle de modèle de réussite, et du rapport virtuel des agents virtuels afin d'identifier les caractéristiques nécessaires pour que nos agents puissent avoir les effets escomptés. L'enchaînement empirique s'est produit en 4 étapes. Premièrement, la construction des agents est passée par la modélisation de leur apparence et de leurs comportements verbaux et non-verbaux afin de créer un modèle computationnel de rapport virtuel (Chapitre 6). Deuxièmement, deux études pilotes perceptives ont permis d'évaluer nos agents virtuels animés auprès d'élèves de 6e et 5e, comparables à ceux de l'étude principale (Chapitre 7). Ces évaluations portaient sur la comparaison de deux types d'agents virtuels (sans *versus* avec comportements de rapport virtuel), ainsi que de l'agent féminin et masculin, sur un ensemble de caractéristiques permettant l'identification (e.g., compétence en mathématiques, sentiment de rapport virtuel, niveau d'anthropomorphisme, sympathie, identification à l'agent et à sa réussite en mathématiques). La troisième étape consistait à élaborer les tests standardisés de mathématiques et à les pré-tester, car les effets de menace du stéréotype s'observent principalement sur les exercices difficiles. Il était important de créer des tests ni trop faciles (effet plafond) ni trop difficiles (effet plancher). Les tests portaient sur les notions importantes, là où les élèves ont le plus de difficulté selon les établissements scolaires (Chapitre 8). Enfin, le dernier chapitre de cette partie est l'étude principale pour évaluer l'efficacité de nos agents virtuels sur la performance en mathématiques des élèves, et plus particulièrement des filles, sur les tests standardisés élaborés au chapitre précédent (Chapitre 9).

Nous avons respecté les bonnes pratiques de la recherche sur l'aspect éthique et sur la science ouverte. Les études ont fait l'objet d'une autorisation du comité éthique d'Aix-Marseille Université, présentée en Annexe B (numéro de dossier 2021-09-30-04), et nous avons demandé l'autorisation parentale de participation à l'expérimentation, puisque les élèves étaient mineurs (Annexe C). Toutes les publications qui ont eu lieu lors de cette thèse sont disponibles sur Hal, et les données sont accessibles sur https://osf.io/zrq2p/?view_only=3a5081da33c043b1adf0e739e2c547bd. L'ensemble des outils développé dans cette thèse est disponible sur https://github.com/ArmandoMarjorie/French_child_IAT. Nous n'avons pas pré-enregistré notre

méthodologie pour s'adapter plus facilement au terrain.

La réalisation de toutes les études dans le cadre de cette thèse reposait sur le volontariat d'un réseau d'établissements issus du pôle pilote Ampiric [1] suite à nos interventions auprès des équipes pédagogiques de mathématiques et de la direction du collège. En effet, toutes les études se sont déroulées directement sur le terrain. Les participants, pour toutes les études, sont donc des échantillons de convenance. Le Tableau 5.1 résume, pour chaque étude de cette partie, l'objectif principal et les caractéristiques de l'échantillon. Au total, 582 élèves ont participé dans cette thèse.

1. https://ampiric.fr/

TABLEAU 5.1

Résumé de chacune de nos études. Les noms des collèges ont été anonymisés.

Étude	Objectif(s)	Taille d'échantillon	Nombre de filles	Nombre de garçons	Collège(s)	Catégorie(s)	Nombre de classes
Étude pilote 0 (Chapitre 6)	Évaluer l'âge de nos agents virtuels à partir de photos.	25	15	10	Collège A	REP+	1 classe de 6e
Étude pilote 1A (Chapitre 7)	Vérifier que nos agents virtuels possèdent tous les prérequis en tant que modèles de réussite, vérifier que les élèves perçoivent le rapport virtuel. Les agents se présentent et n'expliquent pas la leçon de mathématiques.	49	23	26	Collège A	REP+	1 classe de 6e, 1 classe de 5e
Étude pilote 1B (Chapitre 7)	Mêmes objectifs que l'étude pilote 1A. Les agents se présentent et expliquent la leçon de mathématiques.	95	44	51	Collège A	REP+	4 classes de 6e, 4 classes de 5e
Étude pilote 2 (Chapitre 8)	pré-tester les tests standardisés de mathématiques, ajuster la difficulté.	38	18	20	Collège B	REP+	1 classe de 6e, 1 classe de 5e
Étude principale (Chapitre 9)	Évaluer l'efficacité de nos agents virtuels sur la performance en mathématiques des élèves, et plus particulièrement des filles par la réduction des effets de menace du stéréotype.	375	178	197	Collège C	REP+	4 classes de 6e, 3 classes de 5e
					Collège D	non-REP	2 classes de 6e, 2 classes de 5e
					Collège E	REP+	5 classes de 6e, 4 classes de 5e
					Collège F	REP+	3 classes de 5e

6. Modélisation et implémentation des agents virtuels pédagogiques et de l'environnement virtuel d'apprentissage

Sommaire

6.1. Objectifs . . . 134
6.2. Agents virtuels et environnement virtuel . . . 134
6.2.1. Environnement virtuel de la salle de classe . . . 135
6.2.2. Création des premiers modèles 3D des agents virtuels . . . 138
6.2.3. Perception de l'âge des agents virtuels . . . 139
6.2.4. Modèles 3D finaux des agents virtuels . . . 140
6.3. Modélisation computationnel du comportement de rapport . . . 142
6.3.1. Bibliothèque de comportements multimodaux de rapport . . . 143
6.3.1.1. Expressions faciales . . . 143
6.3.1.2. Gestes et stratégies de communication . . . 148
6.3.2. Association des comportements non-verbaux et verbaux . . . 157
6.4. Intégration des agents virtuels dans l'environnement virtuel d'apprentissage interactif . . . 163
6.5. Création de trois types d'agents virtuels . . . 171
6.6. Discussion . . . 171
6.7. Points clés du chapitre . . . 173

6.1. Objectifs

Ce chapitre détaille le procédé de modélisation, de création, et d'implémentation des agents virtuels pédagogiques ainsi que de l'environnement virtuel d'apprentissage. Nous avons deux objectifs dans ce chapitre : (1) créer des agents virtuels incarnant des modèles de réussite pour des enfants en 6e et en 5e, (2) modéliser leur comportements verbaux et non-verbaux de sorte à créer un rapport virtuel. La première section est consacrée à notre premier objectif. Cela consiste à créer des agents virtuels auxquels les élèves puissent s'identifier (Section 2.3.4.3). Plus il y a de similitudes entre l'agent et l'élève, plus l'élève peut s'identifier. L'apparence de l'agent est un critère important puisqu'elle peut induire des informations comme l'âge perçu, les origines culturelles perçues, ou encore le genre perçu. L'apparence peut donc fournir des similitudes et des traits susceptibles d'améliorer l'identification. La majorité des élèves participants à nos études ont entre 10 et 12 ans. Par conséquent, l'apparence de nos agents doit être adaptée à cette population. En particulier, les élèves doivent percevoir les agents comme légèrement plus âgés qu'eux. S'ils paraissent trop âgé, les élèves ne pourront pas s'identifier car ils seront considérés comme trop éloigné. S'ils paraissent trop jeune ou du même âge que les élèves, ils pourraient être découragés et ne pourront pas s'identifier. Nous visons donc à créer des agents virtuels perçus comme étant légèrement plus âgés que les élèves. Nous détaillons également dans la première section la salle de classe virtuelle où les agents virtuels prennent place, avec les différents éléments en 3D.

La deuxième section porte sur la façon dont nous avons modélisé le comportement humain de nos agents, avec des comportements verbaux et non-verbaux permettant de créer un rapport virtuel (GRATCH et al., 2006). Nous avons dû faire face à un certain nombre de contraintes techniques, notamment la grande difficulté à prendre en compte les signaux des utilisateurs puisque toutes nos études se déroulent directement dans les classes. L'objectif de cette section est donc de créer un modèle computationnel de rapport virtuel sans analyser les comportements non-verbaux des élèves. Nous décrivons la bibliothèque de comportements multimodaux que nous avons créé pour cela, ainsi que le modèle computationnel qui en résulte. Enfin, nous terminons ce chapitre par l'implémentation des éléments en 3D et du modèle computationnel de rapport virtuel pour permettre la création de notre application, et nous discutons de la création de ces agents.

6.2. Agents virtuels et environnement virtuel

Cette section est consacrée uniquement aux visuels de l'application, c'est-à-dire à l'environnement virtuel, aux objets en 3D, et aux personnages en 3D qui la composent. L'environnement virtuel est une salle de classe classique où les agents virtuels sont assis sur une chaise. Nous décrivons premièrement la salle de classe, puis les étapes de construction des modèles 3D de nos agents virtuels pour en faire des modèles de

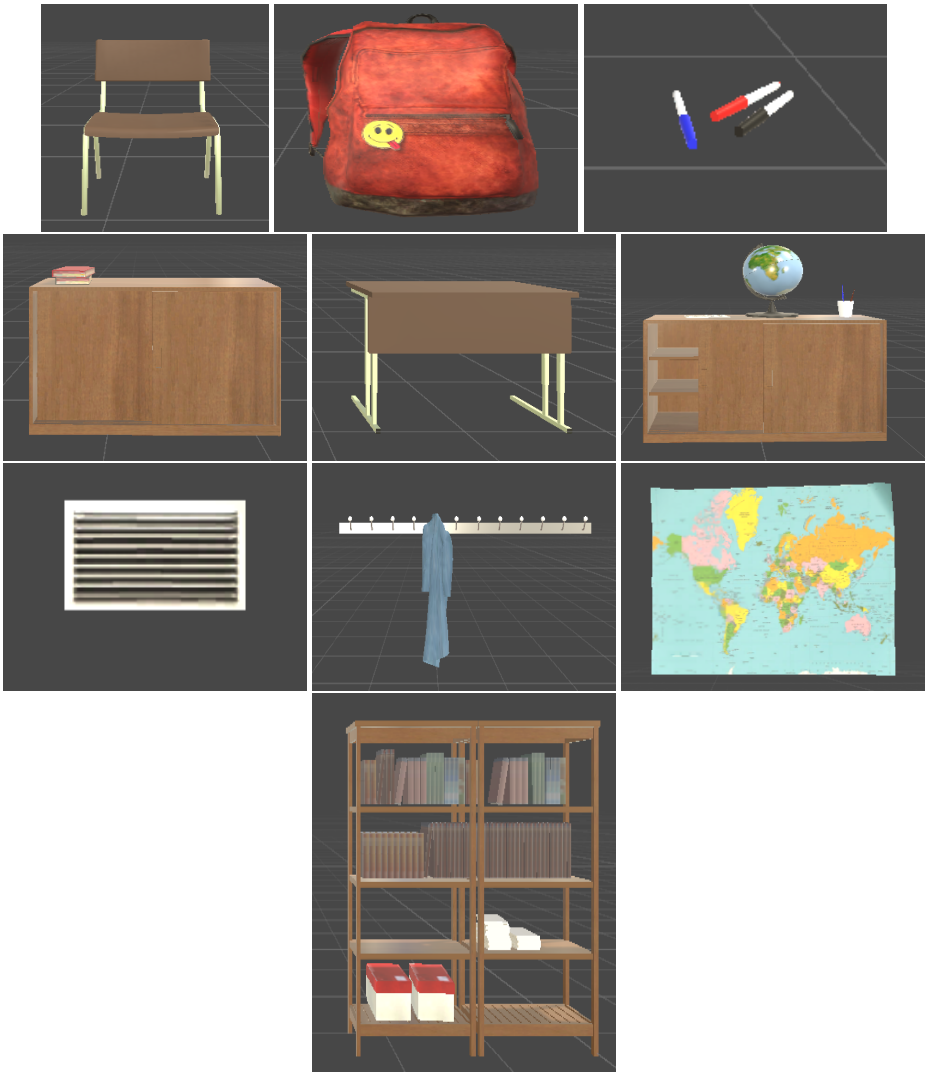
réussite. Les animations sont décrites dans la section suivante.

6.2.1. Environnement virtuel de la salle de classe

L'environnement virtuel a été créé de manière à ne véhiculer aucun stéréotype. Nous avons suivi les résultats des études de CHERYAN, MELTZOFF et KIM (2011) : lorsqu'une salle de classe virtuelle contenait des décorations qui reflétaient des stéréotypes masculins liés aux STIMs, les participantes se sentaient moins à leur place que les participants. De plus, elles pensaient avoir moins de chance de réussir en STIM que leurs collègues masculins à cause de ce sentiment de non-appartenance. Ce n'est que dans une salle de classe virtuelle sans décoration stéréotypique que les participantes et participants ont exprimé les mêmes chance de réussite et le même sentiment d'appartenance. Nous avons donc utilisé une salle de classe virtuelle semblable aux salles de classe françaises habituelles, sans aucun objet stéréotypique lié aux STIMs. La décoration comprenait cinq bureaux et douze chaises de cours, deux étagères avec des livres, des boîtes, et des feuilles de papier, un sac-à-dos posé sur le sol, deux commodes, l'une avec un livre posé par dessus et l'autre avec un globe terrestre et du matériel de peinture posés dessus, et trois stylos posés sur un bureau. La décoration murale comprenait une carte du monde, deux grilles d'aérations, et un porte-manteaux avec une veste accrochée à celui-ci. Tous les éléments sont illustrés en Figure 6.1.

FIGURE 6.1.

Illustrations de tous les éléments de décoration en 3D présents dans la salle de classe virtuelle.



Ces éléments 3D proviennent d'une bibliothèque d'objets pour Unity [1]. Nous avons mis en place ces éléments dans une scène pour en faire une salle de classe, illustrée en Figure 6.2. Nous avons créé deux versions : une application directement sur le web ainsi qu'une application pouvant fonctionner sur tablette. Les établissements participants n'ayant pas tous accès à des ordinateurs compatibles avec Unity Web Player pour la version web, il était nécessaire de leur prêter des tablettes. Plus particulièrement, nous leur avons prêter des Ipads (modèle MR7F2NF/A, iOS 11.4), ce qui a imposé des contraintes d'application plus légère. Nous avons donc placé des éléments 3D sur la scène uniquement s'ils étaient dans l'angle de la caméra (Figure 6.3), pour optimiser les calculs et donc les FPS (*frames per second*, e.g., le nombre d'image affi-

1. https://assetstore.unity.com/packages/3d/props/interior/school-classroom-60343

chée par seconde). Nous avons effectué des tests de faisabilité sur plusieurs dispositifs (les Ipads, nos ordinateurs ainsi que ceux des établissements scolaires), ce qui nous a permis d'optimiser les textures avec un applat de couleur pour un rendu plus optimal, notamment sur tablette.

FIGURE 6.2.

Salle de classe virtuelle.



FIGURE 6.3.

Différents point de vue de la scène 3D de la salle de classe virtuelle sous Unity 3D.



Note. Des éléments sont absents (porte, mur, tableau) car ils ne sont pas dans l'angle de vue de la caméra.

6.2.2. Création des premiers modèles 3D des agents virtuels

Nous avons créé deux agents virtuels adolescents, l'un ayant l'apparence d'une jeune fille, l'autre d'un jeune garçon. Les modèles 3D ont été créé en étroite collaboration avec l'entreprise *Davi the Humanizers* avec le logiciel Character Creator 3. Nous avons commencé par créer l'agent féminin. Sa création s'est déroulée en plusieurs étapes. La première version de l'agent féminin est présentée en Figure 6.4. Pour avoir un premier retour d'enfants, nous avons demandé l'avis de 5 filles de 6e de notre entourage. Plusieurs problèmes ont émergé : l'agent virtuel paraissait trop âgé (une vingtaine d'années), trop "parfait" (peau trop lisse) et trop mince, et sa coupe de cheveux ne convenait pas car peu de filles avait cette coupe. Elles auraient voulu voir un agent avec des cheveux long ou mi-long, ou bien attaché en queue de cheval. Elles ne se sont pas senties représentées par cet agent qui ne leur ressemblait pas.

FIGURE 6.4.

Première version (panel A) et deuxième version (panel B) de l'agent féminin.
Panel A.



Panel B.



Note. Le débardeur est celui par défaut et ne représente pas la tenue finale de l'agent.

En prenant en compte ces considérations, nous avons créé une deuxième version de l'agent féminin (Figure 6.4). Étant en collaboration avec un collège en REP+, nous avons pu comparer ce deuxième agent avec le trombinoscope des élèves de 3e, car

cette population est légèrement plus âgée que les élèves de nos études (6e et 5e). Notre agent ne ressemblant pas aux élèves du trombinoscope, nous avons cette fois-ci créé des agents à partir de photos de vrais adolescents afin qu'ils ressemblent le plus possible à la majorité de ces élèves. Nous ne pouvons pas utiliser les photos du trombinoscope par rapport aux droits d'image. Utiliser une vraie photo pour créer un agent comporte cependant des contraintes. Les photos doivent être libres de droit, être en haute définition, et les personnes sur ces photos doivent aborder une expression faciale le plus neutre possible. Il est cependant difficile de trouver des photos d'adolescents ou d'enfants respectant ces critères. La figure 6.5 montre les photos que nous avons trouvé, ainsi que les agents qui en résultent.

FIGURE 6.5.

Photos libres de droits d'adolescents réels à gauche, dont le visage a inspiré les agents à droite.



6.2.3. Perception de l'âge des agents virtuels

Afin de construire des agents capables d'incarner des modèles de réussite, nous avons évalué deux agents virtuels sur la perception de leur âge, un agent masculin et un agent féminin, auprès d'une classe de 6e d'un collège en REP+. Ces élèves sont comparables à la grande majorité de nos participants de l'étude principale. Nous évaluons uniquement l'âge des agents car il peut être déduit de l'apparence. Une évaluation plus poussée (e.g., identification à l'agent, perception du rapport virtuel) est proposée au prochain chapitre.

Il s'agissait ici d'évaluer des images fixes des deux agents virtuels, masculin et féminin, en ligne. L'enseignante de mathématiques a fourni aux élèves le lien vers le

sondage. Ils ont ainsi pu le compléter chez eux. La Figure 6.6 montre les images que les élèves ont évalué. Un total de 25 élèves (15 filles, 10 garçons, M_{age}=11.08, ES=.06) a estimé un âge pour chaque agent. En moyenne, l'agent masculin a été perçu comme ayant 12.05 ans (ES=.48) et l'agent féminin comme ayant 15.23 ans (ES=.75). Deux élèves ont été retiré de l'évaluation de l'âge de l'agent féminin, l'un ayant estimé son âge à 42 ans, l'autre à 50 ans.

FIGURE 6.6.

Images fixes des agents évalués par une classe de 6e.



6.2.4. Modèles 3D finaux des agents virtuels

Pour incarner un modèle de réussite, les agents doivent être perçus comme étant légèrement plus âgés que les participants, ce qui était le cas pour l'agent féminin, mais pas pour l'agent masculin qui était perçu comme ayant le même âge que les participants. Nous avons donc gardé cet agent féminin, puis nous l'avons masculinisé pour créer un autre agent masculin, en suivant l'étude de NAG et YALÇIN (2020). Ainsi, les yeux de l'agent masculin sont plus petits et moins en amande sur les côtés extérieurs, ses cils sont plus courts, ses sourcils sont plus épais, il a une arcade sourcilière légèrement plus nette et plus marquée, des pommettes plus saillantes, un nez légèrement plus large, un peu de moustache et de barbe, un dessin de la bouche un peu différent (moins vague sur le dessus), une mâchoire un peu plus proéminente, sa couleur des lèvres est différente, ses cheveux sont courts, sa pomme d'Adam est plus visible, et il a un cou et des épaules plus larges. Étant donné l'origine culturelle prédominante des élèves, nous avons donné à nos agents l'apparence d'une jeune

fille maghrébine et d'un jeune garçon maghrébin (couleur de peau plus ou moins foncée, cheveux non lisses, couleurs des yeux foncés). Nous avons attribué les mêmes vêtements avec les mêmes couleurs aux deux agents pour ne pas qu'il y ait une différence de perception entre les deux agents par rapport à cela, et également car un modèle de réussite incarné par une fille "trop" féminine ou "trop" masculine peut ne pas permettre l'identification (CHERYAN, SIY et al., 2011 ; PRONIN et al., 2004). Nous avons opté pour des vêtements unisexes portés par un grand nombre de personnes (un pull, un jean, des baskets). L'apparence des agents finaux est illustrée dans la figure 6.7.

FIGURE 6.7.

Modèles 3D des agents virtuels finaux.



Note. Les couleurs des pulls ont été ensuite changées sous Unity pour avoir les mêmes, en gris.

L'évaluation des agents décrite dans cette section portait uniquement sur la perception de leur âge, basée uniquement sur leur apparence. Il est en effet impossible d'évaluer le niveau d'identification, d'intelligence, ou encore de sympathie d'un agent sur la base de simples photos. Ces évaluations seront traitées au chapitre suivant, avec les agents virtuels animés. Nous avons placé les agents virtuels finaux dans l'environnement virtuel, assis sur une chaise face à la caméra (Figure 6.8) dans une version web et une version locale.

FIGURE 6.8.

Modèles 3D des agents virtuels finaux dans l'environnement virtuel, assis sur une chaise face à la caméra.



6.3. Modélisation computationnel du comportement de rapport

La prochaine étape dans la création de nos agents virtuels est de les animer avec l'objectif de simuler un comportement induisant un rapport virtuel. Plusieurs modèles computationnels de rapport utilisent les signaux des utilisateurs pour s'adapter à eux, notamment en coordonnant les gestes des agents à ceux des utilisateurs. Ces modèles supposent un accès aux signaux non-verbaux de l'utilisateur, via une webcam et des outils de reconnaissance des expressions faciales, gestes ou posture des utilisateurs comme OpenFace (BALTRUSAITIS et al., 2018) ou OpenPose (QIAO et al., 2017). Ils ne conviennent pas dans notre cas étant données les contraintes techniques. En effet, toutes nos études se déroulent hors laboratoire, directement dans les classes où il est difficile de mettre en place ce genre de dispositif. Notre objectif est donc de créer un modèle de rapport sans avoir accès à ces signaux. Le comportement des

agents se caractérisent par une suite d'animation sur une durée d'interaction. Le choix des animations est particulièrement important pour créer un sentiment de rapport virtuel, afin de faire émerger les trois composantes du rapport : la positivité (Section 3.4.1.1), l'attention mutuelle (Section 3.4.1.2), et la coordination (Section 3.4.1.3). Nous présentons les animations ci-dessous.

6.3.1. Bibliothèque de comportements multimodaux de rapport

Nous avons défini des animations grâce aux travaux sur le rapport, dont nous avons fait un état de l'art dans les interactions humain-humain (Section 3.4.1) et agent-humain (Section 3.4.2). Chaque animation est associé à une étiquette spécifique. Le comportement non-verbal comprend les gestes, les postures, et les expressions faciales. Le comportement verbal concerne les stratégies de conversation connues pour favoriser le rapport. Nous avons animé les éléments du visage pour créer des expressions faciales positives, en particulier les sourires et les mouvements des sourcils, que nous décrivons ci-dessous. Nous décrivons ensuite les stratégies de communication favorisant le rapport ci-dessous, ainsi que les gestes et postures souvent associés à ces stratégies.

6.3.1.1. Expressions faciales

D'après N. WANG et GRATCH (2009), les expressions faciales sont les expressions les plus universelles et les plus puissantes des émotions positives. Parmi elles, les sourires permettent de favoriser le rapport en augmentant la composante de la positivité. Cependant, un sourire n'équivaut pas forcément à de la joie. Quelqu'un peut sourire dans une situation négative. Par exemple, un sourire spécifique apparaît dans l'expression faciale de gêne (KELTNER, 1995) ou d'anxiété (HARRIGAN & O'CONNELL, 1996). On peut en effet distinguer différents types de sourire ayant des significations différentes (OCHS et al., 2010). Le plus courant est le sourire de Duchenne, aussi appelé "sourire authentique". Il s'oppose au sourire non-Duchenne, ou "faux sourire" (FRANK et al., 1993). Les différents types de sourire sont souvent distingués lors d'interaction interpersonnelle, mais également dans les interactions avec un agent virtuel (NIEWIADOMSKI & PELACHAUD, 2007). Nous avons créé différents sourires en changeant leur intensité et les différents *blendshapes*, illustrés en Figure 6.9. Les *blendshapes* correspondent à des caractéristiques morphologiques que l'on a modifié en suivant le guide du *Facial Action Coding System* (FACS) (EKMAN & FRIESEN, 1978) d'EKMAN et FRIESEN (1978). Le FACS, largement adopté dans les sciences du comportement, est probablement la méthode prédominante pour encoder les expressions faciales. Il décrit les expressions à travers 46 composants, correspondant approximativement aux mouvements distincts des muscles du visage. D'après EKMAN et FRIESEN (1978), certains AUs (*Action Units*) sont généralement associés à des expressions de joie. Les sourires communiquant de la joie inclus typiquement l'activation du muscle

zygomatique majeur (AU 12), une symétrie de ce muscle, et les joues relevées (AU 6) (EKMAN & FRIESEN, 1978; OCHS et al., 2010; N. WANG & GRATCH, 2009).

Nous avons également créé différentes expressions où les sourcils sont plus ou moins élevés. Nous avons fait lever les sourcils lorsque l'agent met l'accent sur un mot ou une expression, selon la prosodie du discours. Les expressions des sourcils sont illustrées en Figure 6.10. Enfin, les clignements des yeux s'effectuent de manière aléatoire, entre 0.5 et 5 secondes entre chaque clignement.

Le Tableau 6.1 montre toutes les expressions faciales avec leur amplitude, composées de *Blendshapes* avec leur amplitude également. Ces expressions peuvent être combinées, de sorte à pouvoir sourire et cligner des yeux en même temps, par exemple.

FIGURE 6.9.

Illustrations des différents sourires.



Sans expression | Intensité 1 | Intensité 2



Intensité 3 | Intensité 4 | Intensité 5

FIGURE 6.10.

Illustrations des différentes expressions des sourcils.

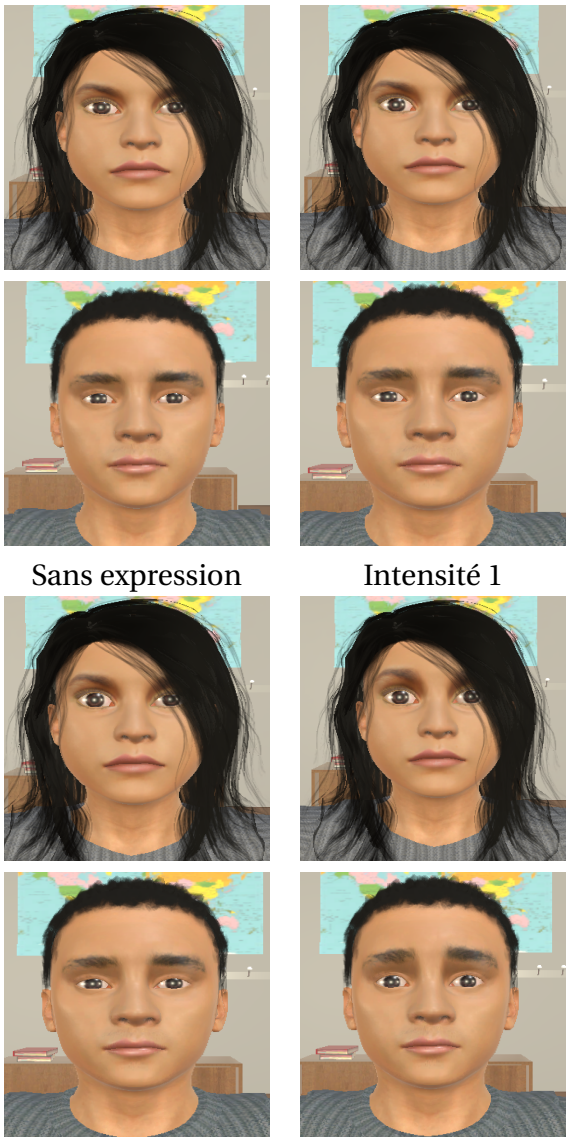


TABLEAU 6.1

Expressions faciales avec leur amplitude (de 0 à 100), leurs Blendshapes, et l'amplitude de chaque Blendshape.

Expression faciale	Amplitude	*Blendshapes*	Amplitude du *Blendshape*
Sourire (intensité 1)	70.5	*brow_raise_inner_L*	45.0
		brow_raise_inner_R	60.0
		eye_squint_L	45.0
		eye_squint_R	45.0
		mouth_smile	60.0
Sourire (intensité 2)	78.2	*mouth_smile*	90.0
Sourire (intensité 3)	100.0	*cheek_raise_L*	50.0
		cheek_raise_R	50.0
		eye_squint_L	35.0
		eye_squint_R	35.0
		mouth_smile	35.0
Sourire (intensité 4)	69.0	*nose_flanks_raise*	10.0
		cheek_raise_L	100.0
		cheek_raise_R	100.0
		eye_squint_L	60.0
		eye_squint_R	60.0
		mouth_smile	80.0
		mouth_smile_L	80.0
		mouth_smile_R	80.0
		brow_raise_inner_L	20.0
		brow_raise_inner_R	20.0
Sourire (intensité 5)	100.0	*mouth_open*	50.8
		cheek_raise_L	52.5
		cheek_raise_R	50.0
		eye_squint_L	60.0
		eye_squint_R	60.0
		mouth_smile	100.0
		mouth_smile_L	30.0
		mouth_smile_R	30.0
Sourcils levés (intensité 1)	100.0	*brow_raise_inner_L*	20.0
		brow_raise_inner_R	20.0
		brow_raise_L	20.0
		brow_raise_R	20.0
Sourcils levés (intensité 2)	57.0	*brow_raise_inner_L*	50.0
		brow_raise_inner_R	50.0
		brow_raise_L	50.0
		brow_raise_R	50.0
Sourcils levés (intensité 3)	100.0	*brow_raise_inner_L*	80.0
		brow_raise_inner_R	80.0
		brow_raise_L	80.0
		brow_raise_R	80.0
		eye_wide_L	60.0
		eye_wide_R	60.0

6.3.1.2. Gestes et stratégies de communication

Nous décrivons les stratégies de communication favorisant le rapport ci-dessous, ainsi que les gestes et postures souvent associés à ces stratégies, comme l'ont fait R. ZHAO et al. (2014) pour leur modèle de rapport. Les modèles de rapport existants dans la littérature sont décrits en section 3.4.2.

Révélation sur soi. Cette stratégie est utilisée pour décrire des préoccupations personnelles. Elle est logiquement caractérisée par une utilisation élevée du pronom personnel "je". L'idée est cependant de rester humble, voire vulnérable. Concernant le comportement non-verbal, on retrouve des hochements de la tête, et le regard doit être porté vers l'élève (MATSUYAMA et al., 2016). Cette stratégie est associée à deux composantes du rapport : l'attention mutuelle et la coordination si l'agent révèle quelque chose de personnel après que l'élève l'ai fait. Notre agent donne son nom après que l'élève lui ai donné le sien (Figure 6.11), coordonnant ainsi la révélation de l'élève. Concrètement, dans la bibliothèque d'animation, nous avons créer des hochements de tête ainsi que des gestes portés vers soi-même (voir Figure 6.12) lorsque l'agent donne des informations personnelles.

FIGURE 6.11.

L'agent demande à l'élève : « C'est quoi ton nom ? ». L'élève utilise le clavier virtuel sur tablette, et le clavier sur ordinateur.



FIGURE 6.12.

Différents gestes portés vers soi-même.



Révélation sur soi (plus intime). Tout comme la stratégie précédente, il s'agit de décrire des préoccupations personnelles, mais cette fois-ci plus intimes. Le partage d'informations privées avec d'autres a démontré des avantages relationnels, également dans les interactions humain-machine (POTDEVIN et al., 2018). Dans notre contexte, nous avons donc défini des phrases spécifiques pour que les agents parlent de leurs bonnes notes en mathématiques et expliquent comment ils les ont obtenu, par exemple : « *j'ai des bonnes notes en maths, parce que j'ai passé beaucoup de temps à réviser et à faire mes devoirs. Du coup j'ai mieux compris les leçons et les exercices. Y'a pas de secret en fait, on me l'avait dit, mais c'est vraiment vrai, faut réviser et écouter en cours, et après ça marche.* ».

Révélation sur soi (négative). Il s'agit ici de faire une révélation plutôt négative sur soi-même. Ceci permet de valider la révélation personnelle négative de l'élève afin qu'il ne se sente pas déconnecté de l'agent. Cette stratégie est donc reliée à la coordination, mais également à la positivité (R. ZHAO et al., 2014). Lorsque l'agent demande à l'élève s'il s'en sort bien en maths pour apprendre à le connaître, l'agent s'adapte à la réponse de l'élève d'une part pour valider la composante de coordination du rapport, et d'autre part pour ne pas rendre l'agent menaçant quant à ses capacités en maths. En effet, si l'élève répond qu'il ne s'en sort pas en maths, l'agent lui répond qu'il ne s'en sortait pas non plus mais depuis qu'il a travaillé, révisé, et fais ses devoirs, ses notes ont augmenté petit à petit, d'où sa réussite actuelle (« *Oui je te comprends, parce que moi aussi ça dépendait des fois. Des fois, je me plantais totalement et des fois ça allait. Mais maintenant, j'ai des bonnes notes en maths, parce que j'ai passé beaucoup de temps à réviser et à faire mes devoirs. Du coup j'ai mieux compris les leçons et les exercices.* »).

Encouragements. D'après KYRIACOU (2010), le rapport dans l'éducation inclue le fait de se soucier des progrès de l'apprenant et de les respecter en tant qu'individu. Ce respect peut être manifesté par des encouragements. L'agent virtuel encourage l'élève à suivre le cours (« *Allez, on va revoir le cours ensemble, tu verras, ça va être bien.* »), et l'encourage à la fin pour qu'il fasse de son mieux lors du contrôle de mathématiques (« *Je suis sûre que tu vas bien réussir.* »).

Référence à une expérience partagée par l'utilisateur. On active ici les "souvenirs" de l'agent virtuel pour que l'élève se sente écouté, et pour montrer que l'agent est attentif à l'élève, ce qui permet d'accroître le rapport pour une meilleure interaction. Cette stratégie permet d'augmenter la notion d'attention mutuelle. L'agent demande à l'élève s'il s'en sort bien en mathématiques (Figure 6.13), adapte sa réponse en fonction de celle de l'élève en utilisant la stratégie de révélation sur soi (négative ou non), puis se rappelle de la réponse de l'élève à la fin de leur interaction pour mieux l'encourager (voir exemple de dialogue en Figure 6.14).

FIGURE 6.13.

L'agent demande à l'élève : « Ça va ? Tu t'en sors bien en maths ? ». L'élève doit cliquer sur une des quatre réponses proposées (« oui ! », « Des fois oui, des fois non. », « Bof. », « Non ! »). L'agent s'adapte alors à la réponse de l'élève.



FIGURE 6.14.

Exemple de dialogues où l'agent fait référence à une expérience partagée par l'élève.
« Voilà, je vais te laisser faire le test de maths. »
[Le discours change selon ce que l'élève avait répondu à la question « Ça va ? Tu t'en sors bien en maths ? » précédemment.]

- **[Choix 1 : « Oui ! »]** « Je suis sûr que tu vas y arriver ! »
- **[Choix 2 : « Des fois oui, des fois non. »]** « Même si tu m'as dit que des fois, ça allait pas trop en maths, je suis sûr que tu vas y arriver ! »
- **[Choix 3 : « Bof. »]** « Même si tu m'as dit que tu t'en sortais pas trop en maths, je suis sûr que tu vas y arriver ! »
- **[Choix 4 : « Non ! »]** « Même si tu m'as dit que tu t'en sortais pas en maths, je suis sûr que tu vas y arriver ! »

Transmission de l'intimité et de la non-dominance. Le concept d'intimité est souvent considéré comme « étant au cœur de nos échanges les plus gratifiants et les plus satisfaisants » (LOMANOWSKA & GUITTON, 2016). Ainsi, si l'intimité a été principalement explorée dans le contexte des relations amoureuses, toutes les interactions peuvent être envisagées à travers ce concept. Dans notre cadre spécifique, l'intimité est considérée comme une compétence sociale dont l'agent virtuel bénéficie pour établir des liens positifs et gratifiants avec l'élève. Nous ne considérons pas l'intimité dans son sens commun, souvent romantique, mais plutôt comme un processus de communication émotionnelle entre deux interactants qui permet de créer des interactions sincères et authentiques, favorisant un sentiment positif et une compréhension mutuelle (REIS et al., 2018). Ces aspects correspondent ainsi aux composantes de la positivité et de l'attention mutuelle du rapport. Les gestuelles intimes ont pour

but d'accentuer l'engagement et la conviction de l'agent virtuel, et de renforcer la dimension sociale de l'interaction (POTDEVIN, 2020). D'après la littérature sur l'intimité dans les interactions entre humains (ARGYLE & DEAN, 1965 ; REIS et al., 2018) et entre humain-agent (OCHS et al., 2015 ; RAVENET et al., 2013), les gestes doivent être plus amples et ouverts (Figure 6.15), ou bien orientés vers soi-même (Figure 6.12).

FIGURE 6.15.

Illustrations de différents gestes amples et ouverts.



Back-channeling. Le *back-channeling* survient pendant une interaction du côté de l'auditeur pour exprimer de l'intérêt, de la compréhension, de la sympathie ou de l'accord. Il s'agit d'un comportement d'écoute, et n'est pas utilisé dans notre cas pour communiquer d'autres informations. L'agent utilise des comportements d'écoute non-verbaux lorsqu'il attend une réponse de l'élève après lui avoir posé une question (Figures 6.11 et 6.13) en hochant la tête plusieurs fois et en souriant, comme dans l'étude de HUANG et al. (2011) avec une séquence de hochements de tête avec une amplitude et une vitesse décroissante.

Transmission de l'empathie. L'empathie affective fait référence au partage des expériences émotionnelles d'autres individus (RENIERS et al., 2011). TETTEGAH et ANDERSON (2007) ont définit l'empathie des enseignants comme une aptitude à communiquer

avec les préoccupations des apprenants, à comprendre leurs préoccupations et à percevoir les situations du point de vue des apprenants. L'empathie affective fait partie intégrante du rapport, et plus particulièrement de l'attention mutuelle. ZHOU (2021) a montré que l'établissement du rapport et de l'empathie dans des contextes éducatifs nécessite le dévouement des enseignants au bien-être des apprenants, et la possibilité aux apprenants de partager et d'exprimer facilement leurs émotions. Notre agent permet cela en demandant à l'élève s'il s'en sort bien en mathématiques, et fait preuve d'empathie affective si l'élève lui fait part de ses mauvais résultats en mathématiques, arbore un sourire d'intensité 1 (Figure 6.9) et des gestes portés vers soi-même (Figure 6.12).

Transmission de l'enthousiasme. Parler avec enthousiasme permet d'augmenter la composante de la positivité, pour ne pas que l'élève se désengage de la tâche, notamment lorsque l'agent explique la leçon de mathématiques. Cette stratégie apparaît alors comme essentielle lors de l'enseignement. Dans notre cas, l'agent virtuel encourage l'élève (« *Allez, on va revoir le cours ensemble, tu verras, ça va être bien.* », « *je suis sûr que tu vas bien réussir.* »).

Enseignement. Que l'on veuille induire le rapport ou non, enseigner requiert des gestes pédagogiques pour transmettre de l'information (CAUNA et al., 2023). L'agent utilise des gestes déictiques pour montrer ce dont il parle (Figure 6.16), et des gestes iconiques pour expliquer un concept (Figures 6.17, 6.18, 6.19, 6.20, 6.21). Pour enseigner avec plus d'enthousiasme, les agents lèvent légèrement le bras avec l'index tendu pour souligner un point du cours présenté verbalement (Figure 6.22). Pour transmettre du rapport lors du cours, l'agent s'adresse directement à l'élève. Par exemple, lorsque l'agent avec le comportement de rapport virtuel montre un exemple à l'élève, il dit « *tu peux voir dans cet exemple* », alors que l'agent sans rapport virtuel dit « *dans cet exemple, on peut voir* ». Lors des exemples, l'agent ne regarde plus l'élève mais l'exemple en question pour diriger son attention vers celui-ci (R. ZHAO et al., 2014).

FIGURE 6.16.

Tous les gestes déictiques.

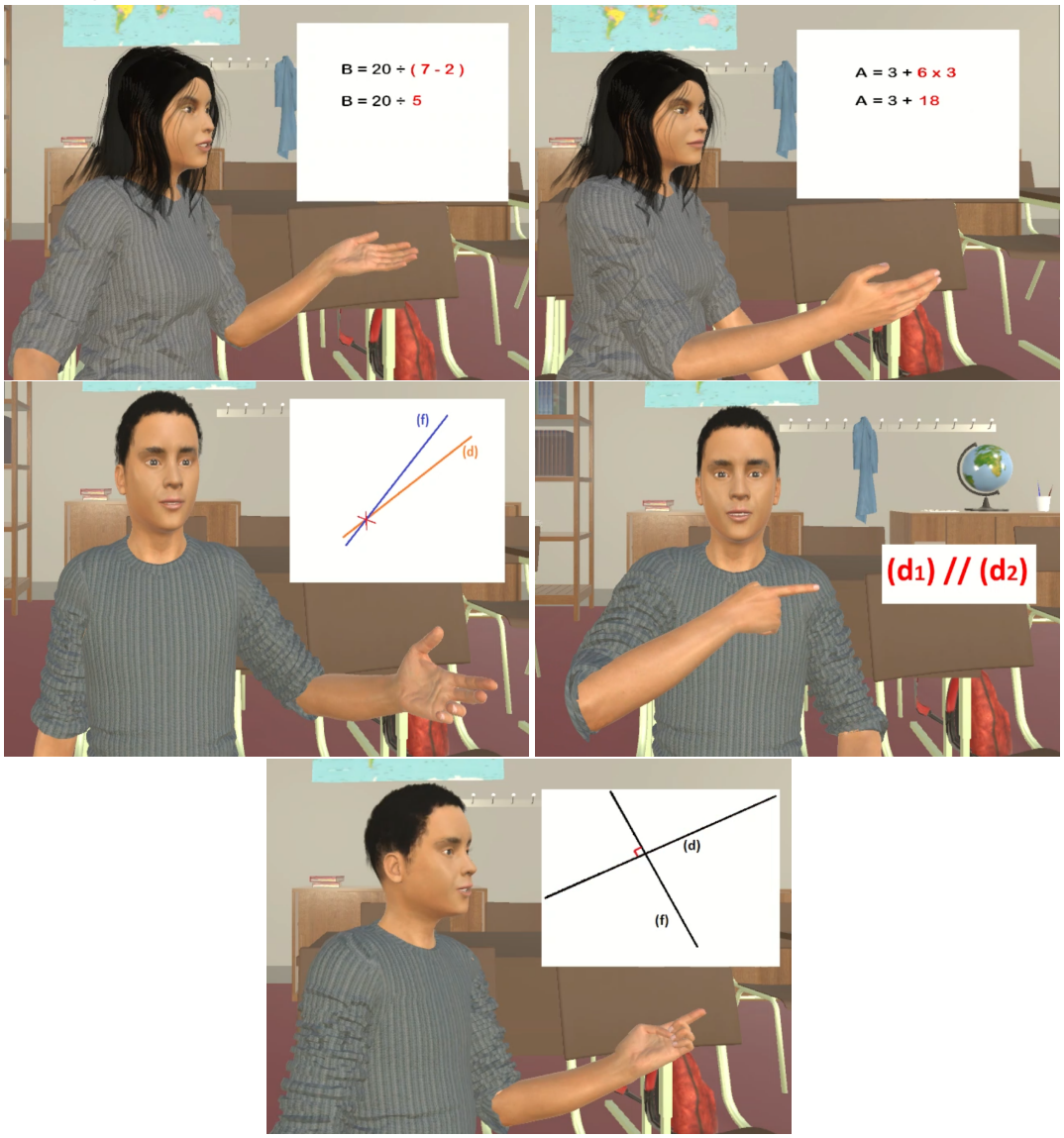


FIGURE 6.17.

Animation du geste iconique de l'addition.



FIGURE 6.18.

Animation du geste iconique de la soustraction.



FIGURE 6.19.

Animation du geste iconique de la division.



FIGURE 6.20.

Animation du geste iconique de la multiplication.



FIGURE 6.21.

Gestes iconiques des droites sécantes (a), des droites parallèles (b), et des droites perpendiculaires (c).

(a) (b) (c)



FIGURE 6.22.

Bras légèrement levé avec l'index tendu pour souligner un point du cours présenté verbalement.



6.3.2. Association des comportements non-verbaux et verbaux

C'est en associant des comportements non-verbaux et verbaux décrits ci-dessus que l'on a créé une animation pour créer le rapport virtuel. Notre modèle computationnel du rapport est représenté par un scénario scripté, où chaque partie du discours représente une animation. L'interaction commence par la présentation de l'agent, où il demande le prénom de l'élève, donne le sien, se présente comme un/une lycéen/lycéenne qui est là pour aider l'élève, et s'intéresse à lui en lui demandant s'il s'en sort bien en mathématiques. L'agent adapte alors sa réponse à celle de l'élève. Une animation est décrite de la façon suivante :

Animation = [stratégie de communication] <étiquette d'animation> discours </étiquette d'animation>

[stratégie de communication] fait référence aux différentes stratégies énoncées dans la section ci-dessus, à savoir faire des révélations sur soi, faire des révélations plus intimes sur soi, faire des révélations négatives sur soi, encourager, faire référence à une expérience partagée précédemment par l'élève, transmettre de l'intimité et de la non-dominance, avoir des comportements de *back-channeling*, transmettre de l'empathie, transmettre de l'enthousiasme, et enseigner. L'étiquette de l'animation correspond à un clip sur Unity. Elle peut correspondre à une expression faciale (e.g., un sourire d'intensité 3), une posture (e.g., bouger légèrement sur la chaise), ou encore un geste (e.g., saluer avec la main). Concrètement, nous avons 8 expressions faciales et 51 gestes. Les Figures 6.23 et 6.24 montre la présentation de l'agent virtuel. Le Tableau 6.2 détaille les étiquettes d'animations utilisées dans les Figures 6.23 et 6.24 afin de mieux comprendre les expressions, gestes et postures de l'agent.

FIGURE 6.23.

Présentation de l'agent virtuel avec des comportements de rapport virtuel.
• [Enthousiasme] <big_smile> <med_brows_up> <tilt> <open_wave> Bonjour! </open_wave> </tilt> </big_smile> </med_brows_up>
• [Intimité] <nod> Et Bienvenue </nod> <small_smile> dans ton cours de <small_brows_up> maths </small_brows_up > </small_smile> <big_smile> sur ordinateur. </big_smile>
• [Révélation sur soi] <big_smile> <hands_on_heart> <nod> <small_brows_up> Je </small_brows_up> suis là </nod> </hands_on_heart> </big_smile> <small_smile> pour t'aider à revoir <med_brows_up> ta leçon </med_brows_up> sur le parallélisme, les priorités opératoires. </small_smile>
• [Intimité] <small_smile> <tilt> <med_brows_up> C'est quoi ton nom? </med_brows_up> </tilt> </small_smile>
• [Back-channeling] <big_smile> <continuous_nod> </continuous_nod> <nod> </nod> </big_smile>
• [Enthousiasme] <smile_teeth> <small_brows_up> Sympa </small_brows_up> <continuous_nod> ton prénom! </continuous_nod> </smile_teeth>
• [Révélation sur soi] <med_smile> <shift> <hand_on_heart> Moi <small_brows_up> c'est {Nora, Dylan}. </small_brows_up> </hand_on_heart> </shift> </med_smile>
• [Révélation sur soi] <med_smile> Je suis <small_brows_up> au <tilt>lycée</tilt> </small_brows_up>, <tilt> et comme toi </tilt>, <open_talking_gesture> j'ai appris les maths au collège. </open_talking_gesture> </med_smile>
• [Intimité] <med_smile> <med_brows_up> Ça va? </med_brows_up> <tilt> Tu t'en sors bien en maths? </tilt> </med_smile>
• [back-channeling] <big_smile> <continuous_nod> </continuous_nod> <nod> </nod> </big_smile>
→ Réponse de l'élève : "*Oui!*" :
- [Enthousiasme] <big_smile> <med_brows_up> Ah </med_brows_up> <small_brows_up> <tilt> oui? </tilt> </small_brows_up> </big_smile> <med_brows_up> Super! </med_brows_up>
- [Révélation sur soi (intime)] <shift> <hand_on_heart> Moi aussi </hand_on_heart> je suis comme toi </shift>

→ Réponse de l'élève : "*Des fois oui, des fois non.*" :
- [Révélation sur soi (négative)] <soft_smile> </soft_smile> <med_smile> <med_brows_up> <tilt> Oui </med_brows_up> </med_smile> je te comprends </tilt>, <hand_on_heart> parce que <med_brows_up> moi aussi </med_brows_up> </hand_on_heart>, ça dépendait des fois. <tilt> Des fois </tilt>, je me <small_brows_up> trompais </small_brows_up> totalement, <shift> et des fois ça allait. Mais maintenant </shift>,

→ Réponse de l'élève : "*Bof.*" :
- [Révélation sur soi (négative)] <soft_smile> </soft_smile> <med_smile> <med_brows_up> <tilt> Oui </med_brows_up> </med_smile> je te comprends </tilt>, <hand_on_heart> <med_brows_up> moi aussi </med_brows_up> au début j'étais <tilt> comme toi </tilt> </hand_on_heart>, <shift> c'était bof. Mais <small_brows_up> maintenant </small_brows_up> </shift>,

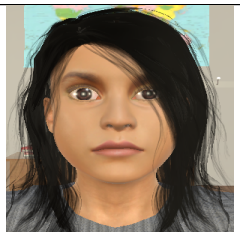
→ Réponse de l'élève : "*Non!*" :
- [Révélation sur soi (négative)] <soft_smile> </soft_smile> <med_smile> <med_brows_up> <tilt> T'inquiète </tilt> </med_brows_up> </med_smile>, <hand_on_heart> moi aussi </hand_on_heart> j'avais du <med_brows_up> mal </med_brows_up> avec les maths et je trouvais ça <tilt> compliqué </tilt>. J'avais des <small_brows_up> notes </small_brows_up> complètement <shift> catastrophiques. Mais maintenant </shift>,

FIGURE 6.24.

Suite de la présentation de l'agent virtuel avec des comportements de rapport virtuel.
• [Révélation sur soi (intime)] < med_smile> J'ai des bonnes notes en <med_brows_up> maths </med_brows_up> parce que <continuous_nod> <hands_on_heart> j'ai passé <med_brows_up> beaucoup </med_brows_up> de temps à réviser </continuous_nod> et à <small_brows_up> faire </small_brows_up> <open_talking_gesture> mes devoirs. Du coup, j'ai mieux compris les leçons et les exercices. </open_talking_gesture> </ med_smile> <big_smile> </big_smile>
• [Enthousiasme] <med_smile> Y'a pas de secret en <med_brows_up> fait </med_brows_up>, <hand_on_heart> <open_talking_gesture> on me l'avait <med_brows_up> dit </med_brows_up> </open_talking_gesture> </hand_on_heart> mais c'est vraiment vrai, faut <hard_nod> réviser </hard_nod> et écouter en <med_brows_up> cours </med_brows_up>, et après ça marche. </med_smile> <big_smile> </big_smile>
• [Encouragements] <med_smile> <gesture_lets_go_arm> <med_brows_up> Allez </med_brows_up>, on va revoir <small_brows_up> le cours </small_brows_up> ensemble </gesture_lets_go_arm>, tu verras, ça va <small_brows_up> être </small_brows_up> bien. </med_smile> <big_smile> </big_smile>
• [Intimité] <med_smile> <continuous_nod> Faudra que tu passes un <med_brows_up> test </med_brows_up> de maths à la fin. </continuous_nod> </med_smile> <big_smile> </big_smile>
• [Encouragements] <med_smile> <hard_nod> Je suis sûr </hard_nod> que tu vas <med_brows_up> bien </med_brows_up> réussir. </med_smile>

TABLEAU 6.2

Description des étiquettes d'animations utilisées dans les Figures 6.23 et 6.24.

Étiquette	Description	Illustration
big_smile	Sourire d'intensité 4.	
med_brows_up	Sourcils levés d'intensité 2.	
tilt	Tête légèrement penchée sur le côté.	
open_wave	Main agitée pour saluer.	
nod	Léger hochement de tête.	
small_smile	Sourire d'intensité 2.	

small_brows_up	Sourcils levés d'intensité 1.	
hands_on_heart	Deux mains posées sur le cœur.	
continuous_nod	Séquence de hochements de tête avec une amplitude et une vitesse décroissante.	
smile_teeth	Sourire d'intensité 5.	
med_smile	Sourire d'intensité 3.	
shift	Repositionnement sur la chaise. L'agent bouge légèrement.	

open_talking_gesture	Bras et mains vers l'utilisateur en parlant. Gestes ouverts.	
soft_smile	Sourire d'intensité 1.	
hand_on_heart	Une seule main posée sur le cœur.	
hard_nod	Hochement de tête prononcé.	
gesture_lets_go_arm	Bras et main qui bougent de l'utilisateur vers l'agent pour lui signifier "allez, viens".	

Nous avons simulé des comportements de rapport en s'appuyant sur la littérature sur le rapport dans les interactions interhumaines et humain-agent, et en créant des scripts fondés sur des balises de type XML pour simuler les différentes stratégies et comportements non-verbaux.

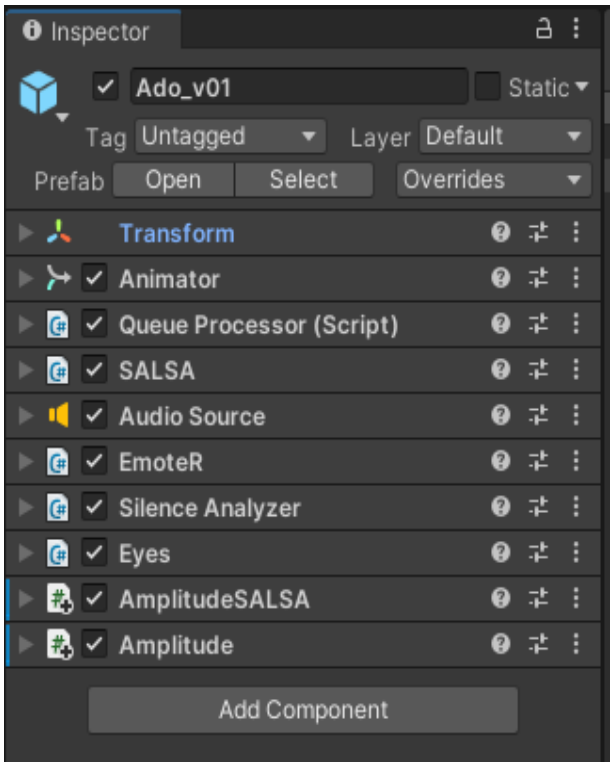
6.4. Intégration des agents virtuels dans l'environnement virtuel d'apprentissage interactif

Nous avons créé deux versions : une application directement sur le web avec Unity Web Player basé sur WebGL, ainsi qu'une application pouvant fonctionner sur tablette avec Unity 3D. Les établissements participants n'ayant pas tous accès à des ordinateurs compatibles avec Unity Web Player, il était nécessaire de leur prêter des tablettes. Plus particulièrement, nous leur avons prêté des Ipads (modèle MR7F2NF/A, iOS 11.4). Tous les éléments ont été mis en place sur Unity3D avec le rendu URP (*Universal Render Pipeline*).

Les agents ont été inclus dans la scène sous Unity en tant que *GameObject*, assis sur une chaise face à la caméra (Figure 6.8). Ils possèdent le module SALSA LipSync v2 pour intégrer la voix et le discours, synchroniser l'animation labiale, créer et utiliser les animations faciales, ainsi que le module AmplitudeSALSA qui associe les modules SALSA LipSync v2 et Amplitude pour permettre la compatibilité avec les plateformes WebGL pour mettre nos applications en ligne, sur un site web que nous avons créé. Les modules associés aux agents sont illustrés en Figure 6.25. Concernant leurs voix, il nous fallait des voix synthétiques d'enfants en français, afin de générer le discours des agents à partir du texte. Leurs voix ont été créées via Acapela-Box, un site proposant un système de text-to-speech ainsi que des voix synthétiques françaises d'enfants payantes. Nous avons acheté les voix Elise et Valentin. Le discours de l'agent a été entré sur Acapela-Box, avec des balises SSML (*Speech Synthesis Markup Language*) pour modifier la prosodie. Les dialogues de l'agent avec rapport contenant les balises SSML sont données en annexe D, et ceux de l'agent sans rapport sont donnés en Annexe E. Les balises SSML ne sont pas différentes selon l'agent. Il en résulte des fichiers audios intégrés aux module SALSA LipSync v2 et AmplitudeSALSA.

FIGURE 6.25.

Modules des agents virtuels.



Nous avons créé un système de tchat limité pour que les élèves puissent interagir avec les agents avec rapport virtuel. Ces agents s'intéressent aux élèves en leur posant des questions. Ils peuvent répondre à l'aide de réponse prédéfinie (Figures 6.13 et 6.14). Nous avons fait ce choix car toutes nos études se passent directement dans les salles de classe. Il est donc impossible que les élèves puissent parler à voix haute. Nous n'avons pas implémenté de tchat textuel, où les élèves pourraient écrire ce qu'ils veulent, pour ne pas les distraire et pour ne pas risquer qu'ils n'aient plus le temps d'effectuer les autres tâches. En effet, les élèves doivent accomplir plusieurs tests et répondre à des questions durant l'étude principale.

Nous avons nous-même créé 8 animations faciales, 41 gestes, et importé 10 gestes de Mixamo [2]. Mixamo est un site web proposant des animations du corps gratuites, cependant, les animations sont souvent plus adaptées pour des jeux vidéo d'action et ne conviennent pas pour notre contexte. Nous avons créé les animations faciales avec SALSA LipSync v2. Nous avons utilisé le module EmoteR de SALSA LipSync v2 qui permet la configuration d'*emotes* pour créer des expressions faciales. Les *emotes* permettent d'utiliser des *Blendshapes* et de changer leur amplitude pour créer une expression faciale. L'animation labiale est synchronisée avec la parole grâce au module SALSA. Ce module ne détecte pas de phonèmes, et s'appuie sur les configurations

2. https://www.mixamo.com/

de *visemes* pour produire des changements d'animation dynamiques adéquats. Un viseme correspond à une forme que prend la bouche lorsque l'agent parle. Par conséquent, l'utilisation de formes de bouche produisant des formes progressivement plus grandes, mais dynamiquement différentes, est essentielle. Tous les *visemes* que nous avons créé sont illustrés en Figure 6.26. Les gestes ont été créé avec l'interface d'animation d'Unity. Chaque membre du squelette peut être animé en lui donnant une amplitude à un instant t. L'animation étant une succession d'images, les images entre deux mouvements d'un même membre sont calculées par interpolation. La Figure 6.27 montre l'interface sous Unity pour l'animation du geste iconique de la division (Figure 6.19).

FIGURE 6.26.

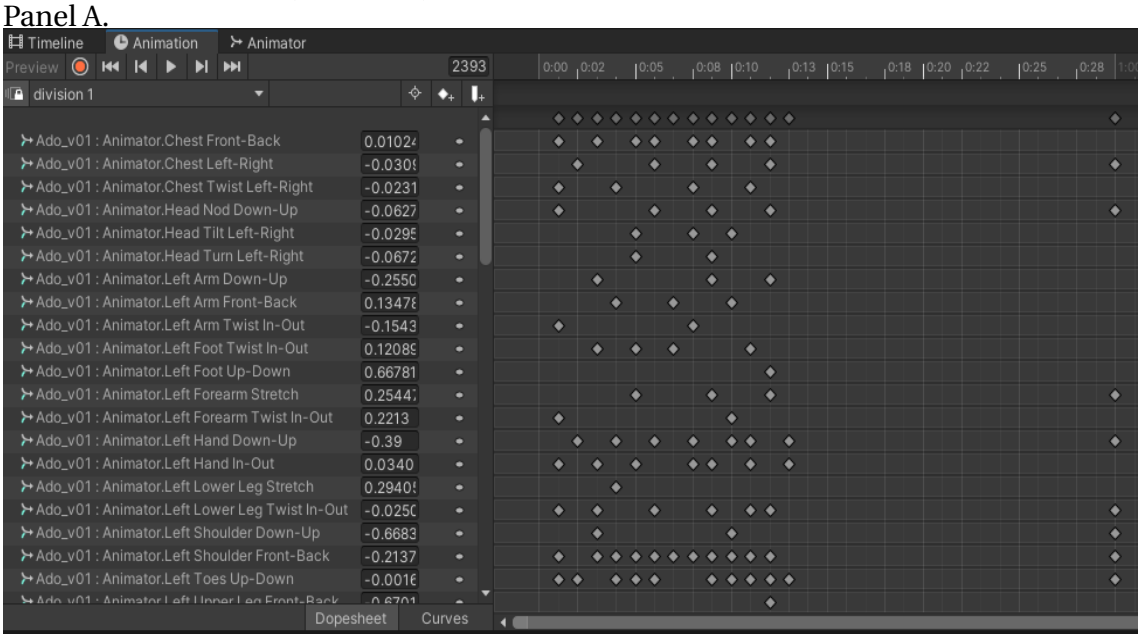
Tous les visemes créés pour l'animation labiale avec le module SALSA.



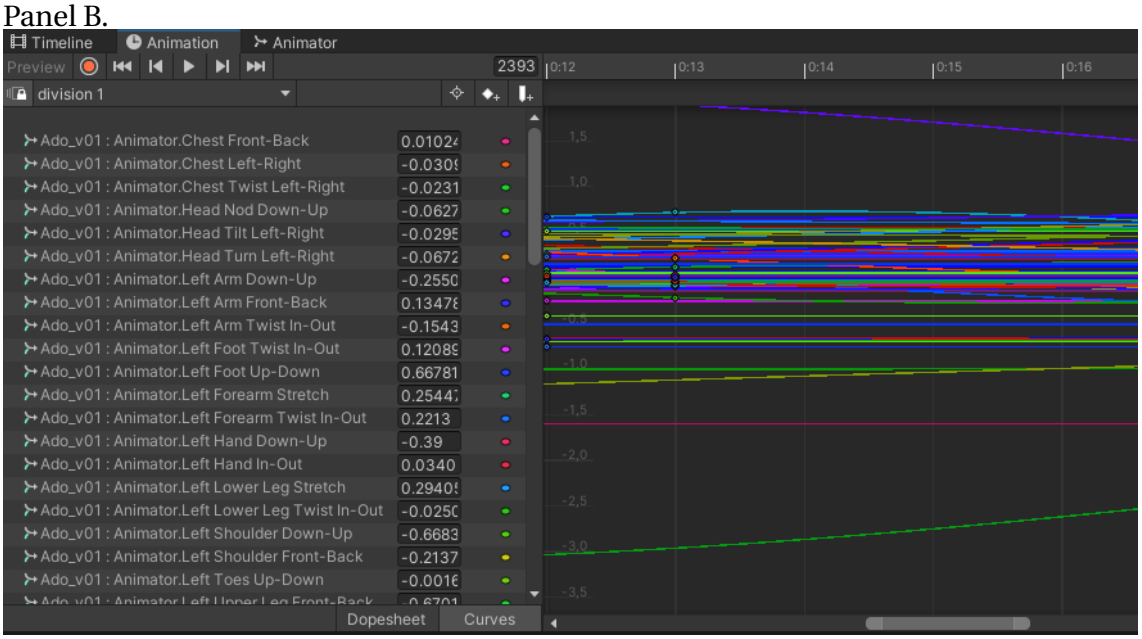
FIGURE 6.27.

Interface d'Unity pour créer une animation, ici le geste iconique pour expliquer la division. Le Panel A montre l'animation représentée par des points, et le Panel B montre la même animation représentée par des courbes.

Panel A.



Panel B.



Pour combiner ces gestes afin de produire des animations, nous avons utilisé un système de masque. Ce système permet de supprimer certaines données d'animation d'un fichier d'animation (e.g., un clip), permettant ainsi au clip d'animer uniquement des parties du personnage plutôt que la totalité. Par exemple, la Figure 6.28 combine

un geste ouvert dont le clip utilise les deux bras, et un geste porté vers soi-même dont le clip utilise également les deux bras. Les masques permettent de combiner ces deux clips pour créer une nouvelle animation. Tous les masques créés et utilisés sont illustrés en Figure 6.29.

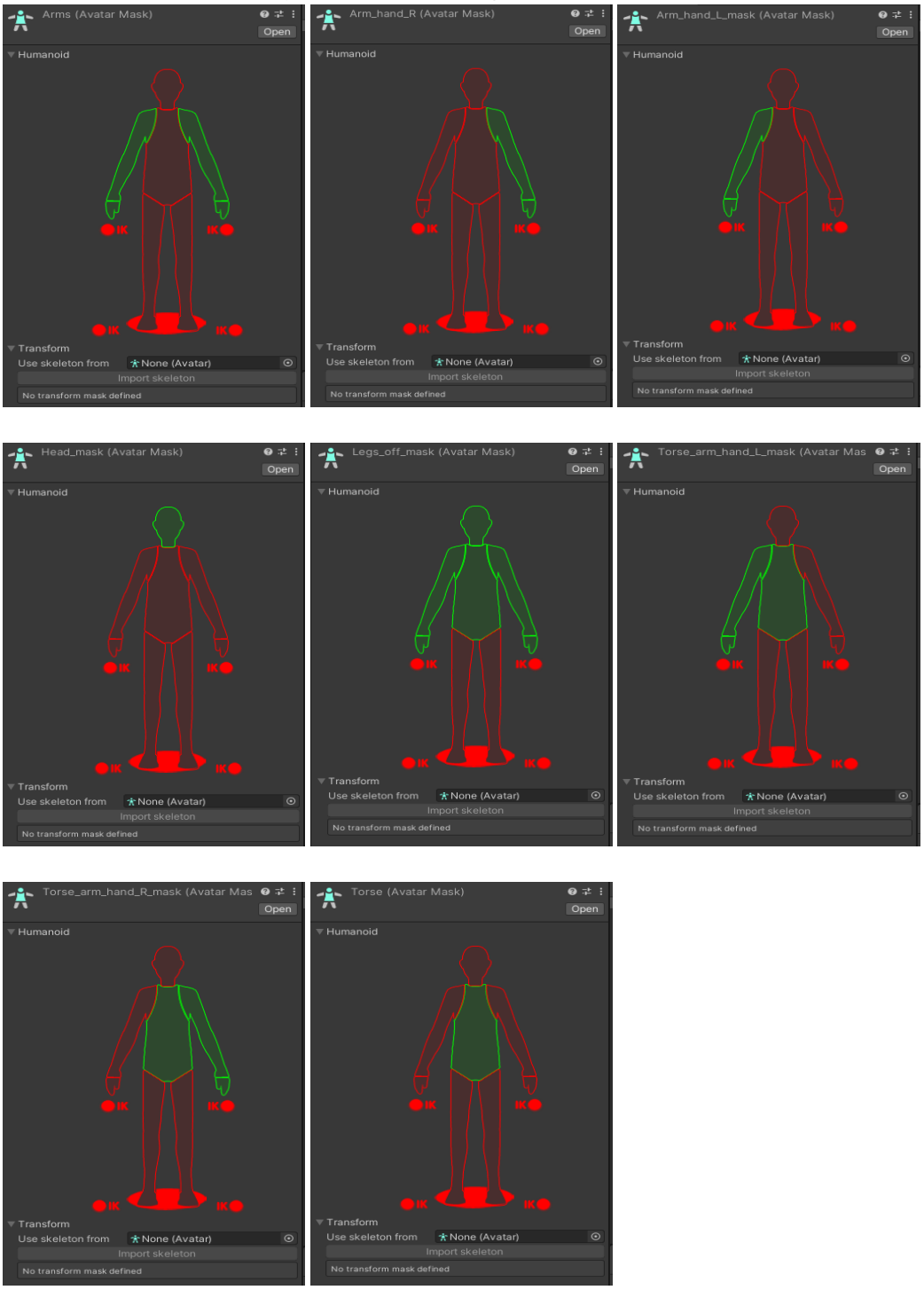
FIGURE 6.28.

Combinaison d'un geste ouvert et d'un geste porté vers soi-même, en utilisant des masques.



FIGURE 6.29.

Masques créés et utilisés pour animer les agents virtuels.



Note. Seules les parties vertes sont animées. Ainsi, si un clip d'animation utilise les bras et la tête et que l'on ajoute le masque *head*, les bras ne bougeront pas.

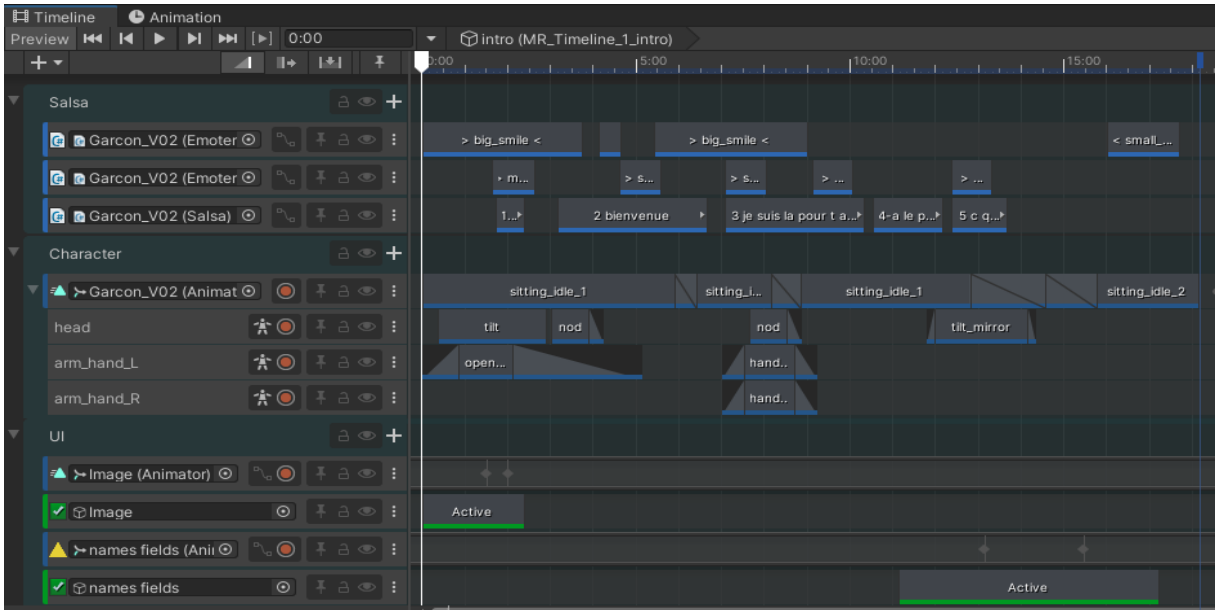
La gestion de l'interaction avec l'agent virtuel se fait via trois *packages* que nous

avons créé : un *package* pour contrôler le discours et les expressions faciales nommé *Salsa* (pour SALSA LipSync v2), un *package* pour contrôler les gestes du personnage ainsi que les masques nommé *Character*, et un autre gérant l'animation de l'interface utilisateur nommé *UI*. Le *package Salsa* comprend un élément pour les mouvements des sourcils, un autre pour les expressions faciales, et un dernier pour le discours. Cette implémentation par *package* modulaire permet la combinaison de plusieurs mouvements, et donc de sourire tout en parlant, par exemple. Le *package Character* contient un seul élément : l'agent virtuel. L'agent virtuel contient alors des masques pour pouvoir animer les différentes parties du corps en même temps. Enfin, le *package UI* contient des animations et des éléments de l'interface utilisateur, notamment le champ pour rentrer son prénom ou encore les réponses prédéfinies des différentes questions posés par l'agent.

La Figure 6.30 montre l'interface sous Unity avec les trois packages comprenant différents éléments. Cet exemple montre l'interaction avec l'agent, du commencement jusqu'au moment où il demande le prénom de l'élève. Dans le *package Character*, l'élément principal (l'agent virtuel *Garcon_v02*, Figure 6.30) possède des animations d'*idles* importés de Mixamo pour qu'il bouge légèrement comme s'il respirait, et également des masques sous forme de sous-éléments (*head*, *arm_hand_L*, et *arm_hand_R*, Figure 6.30). Le nombre de masques varie d'une interaction à une autre. Ainsi, dans une interaction où l'agent adopte uniquement des comportements d'écoute (e.g., lorsqu'il attend une réponse de l'élève), seul le masque *head* est présent, alors que 6 masques (*legs_off, torso_arms_hand_L, torso_arms_hand_R, head, arm_hand_L, arm_hand_R*) sont représentés au moment où l'agent virtuel coordonne sa réponse à celle de l'élève à la question « tu t'en sors bien en maths ? ».

FIGURE 6.30.

Interface sous Unity montrant une interaction avec l'agent virtuel, et les packages pour animer le personnage (Salsa et Character) et pour gérer l'interface utilisateur (UI).



Nous avons codé des scripts en C# avec l'API d'Unity pour le *back-end* de l'application. Ceci concerne la gestion des différentes scènes et interactions, la gestion des réponses des élèves, les événements comme les clicks, et la sauvegarde des données dans des fichiers cachés, seulement accessibles par nous-même. Nous avons enregistré la classe de l'élève, le nom de son établissement, le numéro du test de mathématiques (1 ou 2), les données liées à l'agent virtuel (genre et type d'agent), ses réponses lors du questionnaire d'auto-évaluation en mathématiques, ses données liées au test d'associations implicites (e.g., le temps global pris pour faire le test, le temps de réponse moyen pour chaque *block* de test, le taux de bonnes réponses, le taux de temps de réponse inférieur à 300 ms, si le score est valide, l'ordre de passation des *blocks* congruents et incongruents, les sous-scores, le score), et les réponses de l'élève au test de mathématiques. Le nom et le prénom de l'élève ont été enregistrés puis immédiatement supprimés en lui donnant un identifiant unique. Nous avons mis un place un processus de récupération de ces fichiers JSON de manière anonyme. Nous avons créé un serveur local en PHP en fixant notre propre adresse IP. Il faut donc être connecté avec notre réseau pour les recevoir et lancer le serveur avec notre fichier PHP. il est donc impossible pour une personne extérieure d'y accéder.

6.5. Création de trois types d'agents virtuels

La prochaine étape décrite au chapitre suivant consistait à évaluer les agents virtuels pédagogiques, cette fois-ci animés. En effet, nous devions nous assurer que notre modèle de rapport fonctionnait bien (e.g., que les élèves perçoivent le rapport virtuel), mais également qu'ils remplissaient leur rôle de modèle de réussite. Nous nous sommes également interrogées si le rapport virtuel pouvait augmenter l'identification à nos agents virtuels pédagogiques. Nous avons donc créé, en opposition à l'agent avec des comportements de rapport, des agents avec des comportements plus neutres dont le but était de ne pas véhiculer de sentiment de rapport virtuel (e.g., seulement quelques changements de posture, pas de sourires). Plus particulièrement, nous avons créé trois type d'agents :

- Un agent virtuel avec notre modèle de rapport détaillé dans ce chapitre, et expliquant sa réussite en mathématiques par le travail et les efforts.
- Un agent virtuel donnant les mêmes explications quant à sa réussite en mathématiques, mais sans modèle de rapport virtuel.
- Un agent virtuel ne donnant aucune explication quant à sa réussite en mathématiques, sans modèle de rapport virtuel.

Le discours des agents virtuels sans comportements de rapport virtuel (e.g., les deux derniers agents) est disponible en Annexe E. Les vidéos de tous les agents virtuels sont disponibles sur https://osf.io/zrq2p/?view_only=3a5081da33c043b1adf0e739e2c547bd. Le dernier agent virtuel a été créé pour simuler une condition contrôle lors de notre étude principale sur l'évaluation des agents sur la performance en mathématiques des élèves, et plus particulièrement des filles. Ces agents sans comportements de rapport virtuel exprimaient moins de gestuelles, pas de sourires, et ne s'adaptaient en aucun cas à l'élève. Ces agents ne posaient en effet aucune question à l'élève. Le temps d'interaction était donc un peu moins long.

6.6. Discussion

En vue de notre étude principale pour évaluer l'impact des agents virtuels pédagogiques pour réduire les effets de menace du stéréotype, nous avons créé des agents virtuels pédagogiques en suivant les études en sciences sociales sur les modèles de réussite. En effet, les études présentées au Chapitre 2 montrent que l'identification à un modèle de réussite permet de réduire les effets de menace du stéréotype, dès lors qu'un certain nombre de similitudes est perçu par l'apprenant. L'apparence est donc importante pour véhiculer des similitudes, mais également pour ne pas provoquer des contre-performances. L'agent virtuel doit être perçu comme étant légèrement plus âgé que l'élève, car s'il est perçu comme étant plus jeune ou ayant le même âge que l'élève, celui-ci peut se sentir inférieur s'ils n'ont pas les mêmes résultats en mathématiques (BAGÈS et al., 2016). Par ailleurs, si l'agent est perçu comme trop âgé, l'élève ne pourra

pas s'identifier non plus car le modèle pourrait être perçu comme n'appartenant pas au même groupe social (J. STEELE, 2003). De plus, un modèle de réussite incarné par une fille "trop" féminine ou "trop" masculine peut également ne pas permettre l'identification (CHERYAN, SIY et al., 2011 ; PRONIN et al., 2004). Physiquement, nous avons modélisé deux agents virtuels, une fille et un garçon, en suivant ces études. Nos deux agents portent les mêmes vêtements neutres (un pull gris et un jean), et ressemblent à la majorité des élèves de 4e et 3e du collège avec qui nous avons collaboré grâce au trombinoscope fourni par l'établissement. Nous avons évalué l'âge perçu de ces agents auprès d'élèves de 6e pour nous assurer qu'ils soient jugés comme étant légèrement plus âgés, ce qui était le cas pour l'agent féminin. Nous avons ainsi créé un autre agent masculin, cette fois-ci en modifiant les traits de l'agent féminin en suivant l'étude de NAG et YALÇIN (2020).

Nous avons ensuite créé un modèle de rapport virtuel en suivant les études sur le rapport dans les interactions humain-humain et humain-agent (Chapitre 3). Ce modèle computationnel de rapport virtuel est fondé sur une bibliothèque d'animations en 3D et d'audio que nous avons créée sous Unity. L'association des différents éléments de la bibliothèque (e.g., animations faciales, discours, ou encore gestes) a été faite avec des scripts de type XML. À ce jour et à notre connaissance, c'est la première fois que de tels agents virtuels, qui suit la littérature sur les modèles de réussite et les comportements de rapport, ont été créés en vue de réduire les effets de menace du stéréotype chez les filles en mathématiques.

Nous avons évalué les agents lors de deux études pilotes que nous présentons dans le prochain chapitre. Nous n'évaluons pas l'agent virtuel qui ne donne aucune explication quant à sa réussite, car son modèle de comportement est le même que l'agent sans comportements de rapport.

6.7. Points clés du chapitre

- Création de l'apparence des agents virtuels en suivant les études en sciences sociales sur les modèles de réussite.
- Validation de l'âge perçu des agents auprès d'une classe de 6e de 25 élèves.
- Création d'une bibliothèque de gestes, postures et expressions faciales en 3D favorisant le rapport, en suivant les études sur le rapport dans les interactions humain-humain et humain-agent.
- Modélisation du rapport virtuel avec des scripts XML par des comportements multimodaux en associant une stratégie de communication, des gestes, des postures, des expressions faciales, et un discours dont la prosodie a été ajustée par SSML.
- Création de trois types d'agents pour l'expérimentation : (1) un agent virtuel avec notre modèle de rapport et expliquant sa réussite en mathématiques par le travail et les efforts, (2) Un agent virtuel donnant les mêmes explications quant à sa réussite en mathématiques mais sans modèle de rapport virtuel, (3) un agent virtuel ne donnant aucune explication quant à sa réussite, sans modèle de rapport virtuel.
- Implémentation sur Unity en créant des *packages* modulaires, et avec les modules SALSA LipSync v2 et AmplitudeSALSA.

7. Études pilotes 1A et 1B : évaluation des agents virtuels animés

Sommaire

7.1. Objectifs . . . 175
7.2. Étude pilote 1A : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent . . . 177
7.2.1. Participants . . . 177
7.2.2. Procédure . . . 177
7.2.3. Mesures . . . 178
7.2.4. Résultats . . . 179
7.2.4.1. Analyses descriptives . . . 179
7.2.4.2. Analyses centrales . . . 182
7.2.5. Discussion . . . 185
7.3. Étude pilote 1B : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent et de leurs explications de la leçon de mathématiques . . . 187
7.3.1. Participants . . . 187
7.3.2. Procédure . . . 187
7.3.3. Mesures . . . 187
7.3.4. Résultats . . . 187
7.3.4.1. Analyses descriptives . . . 187
7.3.4.2. Analyses centrales . . . 190
7.3.5. Discussion . . . 195
7.4. Discussion . . . 196
7.5. Points clés du chapitre . . . 199

Ce chapitre est adapté de :

Armando, M., Régner, I., & Ochs, M. (2023). Toward a Tool Against Stereotype Threat in Math : Children's Perceptions of Virtual Role Models. *International Workshop on Multimodal, Interactive Interfaces for Education in International Conference on Multimodal Interaction (ICMI)*, Paris, France. https://doi.org/10.1145/3610661.3616179 (ARMANDO et al., 2023).

7.1. Objectifs

Le chapitre précédent était consacré à la création d'agents virtuels pédagogiques spécifiquement conçus dans le cadre de nos recherches. Ils possèdent un certain nombre de caractéristiques en vue de notre étude principale portant sur l'impact de ces agents virtuels pédagogiques sur la performance en mathématiques des élèves. Plus particulièrement, nous visions à réduire les effets de menace du stéréotype à l'aide de nos agents virtuels incarnant des modèles de réussite. Pour cela, ces agents ont une apparence physique adaptée à des élèves de 11 à 13 ans (niveau 6e et 5e) : ils sont légèrement plus âgés que ces élèves et se présentent en tant que lycéens. Nous avons créé deux agents de genre différent, une fille et un garçon. Notre hypothèse principale était qu'un agent virtuel incarnant un modèle de réussite, capable d'expliquer sa réussite en mathématiques par les efforts et le travail, ainsi que d'instaurer un rapport virtuel pendant l'interaction avec l'élève, pouvait réduire les effets de menace du stéréotype. De plus, nos agents doivent être crédibles, c'est-à-dire suffisamment réalistes, mais pas trop pour ne pas qu'il y ait de décalage entre le réalisme de son apparence et de ses mouvements. Ce décalage, illustré par le phénomène de la *vallée de l'étrange*, peut créer un sentiment de malaise qui entraîne un rejet de l'agent virtuel (KÄTSYRI et al., 2015).

Bien que quelques études aient créé et utilisé des agents virtuels en tant que modèle de réussite en vue d'encourager les femmes en mathématiques (R. ROSENBERG-KIMA et al., 2008 ; R. B. ROSENBERG-KIMA et al., 2010), ces études n'ont pas mesuré si les participants s'identifiaient aux agents. De plus, que ce soit pour ces études ou celles sur le rapport virtuel, la grande majorité de ces travaux a porté sur des agents virtuels adultes avec des participants également adultes. Ceci soulève deux questions : (1) *Des élèves de 6e et 5e peuvent-ils s'identifier à un modèle de réussite incarné par un agent virtuel pédagogique adolescent ?* (2) *Peut-on induire une relation positive via le rapport virtuel entre des élèves de cette tranche d'âge et un agent virtuel adolescent dans une tâche d'apprentissage ?*

L'objectif de ce chapitre est de répondre à ces questions afin de vérifier si nos agents virtuels possèdent tous les prérequis, en tant que modèles de réussite, pour réduire les effets de menace du stéréotype. Nous avons déjà montré que les élèves percevaient bien les agents virtuels comme légèrement plus âgés qu'eux (Chapitre 6). Nous

souhaitons maintenant examiner un certain nombre de critères lorsque les agents sont animés. Tout d'abord, il s'agit de vérifier que tous les agents sont bien perçus comme étant compétents en mathématiques. Cet élément est central pour un modèle de réussite (Chapitre 2). Nous avons effectué un test de Student de comparaison à une norme pour s'assurer que le niveau de compétence perçu de l'agent était très élevé, et donc significativement supérieur au point médian de l'échelle. Nous nous sommes ensuite assurées qu'il n'y avait pas de différence en fonction du genre de l'agent et du genre des élèves. Dans un second temps, nous avons testé si les élèves ont bien perçu le rapport virtuel en comparant les conditions avec et sans rapport virtuel. Nous voulions également explorer si le sentiment de rapport virtuel permettait de mieux évaluer le niveau de compétence de l'agent en mathématiques et de mieux s'y identifier. En d'autres termes, nous cherchions à savoir si le rapport virtuel pouvait faire de nos agents de meilleurs modèles de réussite en mathématiques. En plus d'effectuer des tests de comparaison des moyennes entre les conditions, nous avons effectué des analyses de corrélations entre les variables dépendantes, ainsi que des analyses de régressions. En outre, une mesure de la sympathie des agents virtuels a permis de connaître le degré d'acceptation des agents par les élèves. Enfin, nous avons examiné le niveau d'anthropomorphisme perçu (e.g., à quel point l'agent est proche de l'humain) pour évaluer la crédibilité de nos agents virtuels.

Pour répondre à ces objectifs, deux études pilotes ont été réalisées, lesquelles partagent plusieurs points communs. Tout d'abord, dans les deux études, l'agent s'est présenté comme étant en réussite en mathématiques grâce au travail et aux efforts, et a annoncé que son rôle était d'aider les élèves en mathématiques (cf dialogues détaillés en Annexe D et E, et lien vers les agents animés sur https://osf.io/zrq2p/?view_only=3a5081da33c043b1adf0e739e2c547bd). Nous n'avons pas évalué l'agent virtuel qui ne donne aucune explication quant à sa réussite, car son modèle de comportement est le même que l'agent sans comportements de rapport (voir Section 6.5 pour tous les agents créés). Ensuite, le genre de l'agent était systématiquement manipulé en variable intersujets. Nous avons fait le choix d'affecter de façon randomisée la classe entière (plutôt que l'élève) à un agent féminin *versus* masculin, de telle sorte à ce que tous les élèves d'une même classe voient le même agent. Ceci dans le but d'éviter toute interrogation chez les élèves susceptible d'interférer avec la tâche, ou tout autre biais potentiel lié à la découverte d'agents différents sur leur tablette. En revanche, les deux études pilotes diffèrent sur le traitement inter- *versus* intrasujets de la variable "rapport virtuel de l'agent". Les études dans le domaine des agents virtuels ont pour habitude de manipuler le rapport virtuel en intrasujets afin de maximiser, par effet de contraste, la perception par les utilisateurs des comportements verbaux et non-verbaux de l'agent qui sont supposés induire un rapport virtuel (HUANG et al., 2011). Notre étude pilote 1A propose donc un traitement intrasujets du "rapport virtuel de l'agent" : chaque élève était confronté deux fois de suite au même agent ; l'ordre d'apparition de l'agent avec ou sans rapport virtuel étant contrebalancé au niveau des élèves au sein de chaque classe. Afin d'éviter un sentiment de lassitude lié à la répétition et à la longueur de la scène, le rôle de l'agent a été limité à la séquence de sa

présentation en tant que modèle de réussite en mathématiques. À l'inverse, l'étude 1B repose sur un traitement intersujets du "rapport virtuel de l'agent", qui a permis d'intégrer en sus la séquence où l'agent explique la leçon. L'étude 1B se rapproche ainsi davantage du protocole de l'étude principale, et permet en outre de tester si l'ajout de la leçon renforce les perceptions des élèves vis-à-vis de l'agent.

En résumé, l'étude pilote 1A repose sur un plan expérimental mixte 2 (genre de l'agent : féminin *versus* masculin) X 2 (agent avec *versus* sans rapport virtuel), avec le genre en facteur intersujets et le rapport virtuel en facteur intrasujets. Quant à l'étude pilote 1B, elle repose sur un plan intersujets 2 (genre de l'agent : féminin *versus* masculin) X 2 (agent avec *versus* sans rapport virtuel). Les classes retenues pour les deux études pilotes ont des profils comparables à la majorité des classes des établissements de l'étude principale (REP+).

7.2. Étude pilote 1A : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent

7.2.1. Participants

Deux classes de REP+ (une classe de 6e et une classe de 5e) ont participé à cette étude pilote, soit un total de N = 49 élèves (26 garçons, 23 filles, M_{age} = 11.39, ES=.09). Onze garçons ont interagi avec les agents masculins et 15 garçons avec les agents féminins. Quatorze filles ont interagi avec les agents masculins et 9 filles avec les agents féminins.

7.2.2. Procédure

Les élèves de chaque classe se sont installés dans la salle informatique de leur collège avec leur enseignant de mathématiques et l'expérimentatrice, pendant une heure de cours de mathématiques. Une tablette équipée de casque audio fourni par Ampiric (MIRAULT et al., 2023) a été prêtée à chaque élève. L'expérimentatrice s'est présentée en tant que membre d'une équipe d'intelligence artificielle à la faculté d'informatique de Marseille réalisant une étude dont l'objectif est de créer un personnage virtuel en 3D pour aider les élèves en mathématiques. Elle a également expliqué que les personnages virtuels avaient été créés à partir de vrais élèves de lycée, de leur apparence physique jusqu'à leur histoire personnelle avec les mathématiques. L'expérimentatrice expliquait ensuite aux élèves qu'elle était là pour recueillir leurs avis sur l'agent virtuel créé par l'équipe de recherche, afin de savoir s'il était utilisable en classe. Pour cela, tous les élèves étaient invités à évaluer l'agent virtuel sur plusieurs critères. Il était bien précisé que le même agent serait présenté deux fois,

mais avec des comportements un peu différents, et qu'il fallait évaluer donc deux fois cet agent. L'expérimentatrice insistait sur le caractère totalement anonyme des réponses, précisait qu'il n'y avait ni bonne ni mauvaise réponse, et que l'important était de donner son avis le plus honnêtement possible. Les élèves étaient ensuite invités à mettre leur casque, à répondre à un bref questionnaire démographique (âge, sexe, fréquence de jeux vidéo), puis à démarrer l'interaction avec l'agent d'une durée approximative de 3 minutes. Tous les élèves découvraient donc une première fois l'agent, qui se présentait en adoptant un comportement avec ou sans rapport virtuel, puis complétaient un questionnaire destiné à recueillir leurs perceptions de l'agent sur plusieurs critères (compétence en mathématiques, sentiment de rapport virtuel, niveau d'anthropomorphisme, sympathie, identification à l'agent et identification à sa réussite). Une fois le questionnaire terminé, la seconde partie démarrait dans laquelle le même agent se comportait avec ou sans rapport virtuel (en fonction de la séquence précédente) puis les élèves complétaient à nouveau le même questionnaire. Les élèves pouvaient poser des questions en levant le doigt s'ils ne comprenaient pas un item du questionnaire. Ceux et celles ayant terminé avant la fin de l'heure étaient invités à rester silencieux et à faire leurs devoirs. À la fin de la séance, les élèves étaient remerciés pour leur contribution importante à cette recherche scientifique.

7.2.3. Mesures

Pour l'ensemble des mesures, les élèves étaient invités à donner leurs réponses sur une échelle de Likert en 5 points (de 1 :"pas du tout d'accord" à 5 : "tout à fait d'accord"), en sélectionnant le chiffre qui correspondait le mieux à leurs avis. L'ensemble des items figure en Annexe J.

Compétence en mathématiques de l'agent. Les élèves évaluaient le niveau de compétence de l'agent en mathématiques à l'aide de 3 items formulés spécifiquement pour la présente étude (e.g., "Est-ce que tu penses que le lycéen/la lycéenne en 3D réussit bien en maths ?"). L'alpha de Cronbach indiquant une bonne cohérence interne (α=.74 pour l'agent avec rapport virtuel, α=.81 pour l'agent sans rapport virtuel), un score moyen était calculé pour refléter le niveau perçu de compétence de l'agent en mathématiques (un score élevé traduisant un niveau élevé de compétence perçue en mathématiques).

Sentiment de rapport virtuel. Les élèves répondaient à 16 items adaptés de deux échelles : la *Virtual Intimacy Scale* (VIS) de POTDEVIN (2020), et la *perception de l'interaction* de CUPERMAN et ICKES (2009). Nous avons adapté 12 items de la VIS sur 15 (e.g., "J'ai trouvé que le lycéen/la lycéenne en 3D a voulu m'aider."), et 4 items de la perception de l'interaction sur 27 (e.g., "J'aimerais revoir le lycéen/la lycéenne en 3D une autre fois."). Nous avons sélectionné les items les plus adaptés pour nos études pilotes. Nous avons choisi la VIS car l'intimité est un concept très proche du rapport. Nous avons uniquement conservé les items du questionnaire de la perception de l'interaction pouvant s'appliquer lorsqu'il n'y a pas de véritable conversation.

Cela permet d'évaluer à quel point un individu ressent le rapport virtuel lors de son interaction avec un agent virtuel. Un item inversé (e.g., "Je me suis senti(e) mal à l'aise avec le lycéen/la lycéenne en 3D.") a été recodé. L'alpha de Cronbach indiquait une bonne cohérence interne (α=.87 pour l'agent avec rapport virtuel, α=.86 pour l'agent sans rapport virtuel). Un score élevé indique un fort ressenti du rapport virtuel lors de l'interaction avec l'agent.

Niveau d'anthropomorphisme. Sept items issus de l'échelle de Godspeed de BARTNECK et al. (2009) étaient utilisés pour évaluer à quel point l'agent virtuel était perçu comme se rapprochant de l'humain. Les items renvoyaient à la fois à l'animation (e.g., "J'ai trouvé que le lycéen/la lycéenne en 3D bougeait comme un robot.") et à l'apparence (e.g., "J'ai trouvé que l'apparence du lycéen/de la lycéenne en 3D était naturelle.") de l'agent. Trois items inversés (e.g., "J'ai trouvé que le lycéen/la lycéenne en 3D avait un aspect robotique.") ont été recodés. L'alpha de Cronbach a indiqué une bonne cohérence interne (α=.79 pour l'agent avec rapport virtuel et pour l'agent sans rapport virtuel). Un haut score signifie que les élèves ont perçu l'agent comme plutôt humain.

Sympathie. Cinq items provenant de l'échelle de Godspeed de BARTNECK et al. (2009) ont été utilisés pour évaluer à quel point les élèves ont trouvé l'agent sympathique. Deux items inversés (e.g., "J'ai trouvé que le lycéen/la lycéenne en 3D était désagréable.") ont été recodés. L'alpha de Cronbach a indiqué une bonne cohérence interne (α=.85 pour l'agent avec rapport virtuel, α=.88 pour l'agent sans rapport virtuel).

Identification à l'agent. Les élèves ont répondu à 4 items adaptés de l'échelle de O'BRIEN et al. (2017) (e.g., "Est-ce que tu voudrais avoir plus de point commun avec le lycéen/la lycéenne en 3D?"). Plus le score est élevé, plus les élèves s'identifient à l'agent. L'alpha de Cronbach a indiqué une bonne cohérence interne (α=.72 pour l'agent avec rapport virtuel, α=.81 pour l'agent sans rapport virtuel).

Identification à la réussite de l'agent. Les élèves ont répondu à 4 items formulés spécifiquement pour la présente étude (e.g., "Comme le lycéen/la lycéenne en 3D, j'ai envie de faire de mon mieux en mathématiques."). L'alpha de Cronbach a indiqué une bonne cohérence interne (α=.78 pour l'agent avec rapport virtuel, α=.76 pour l'agent sans rapport virtuel). Un haut score indique que la réussite en mathématiques de l'agent a inspiré les élèves.

7.2.4. Résultats

7.2.4.1. Analyses descriptives

Le Tableau 7.1 présente la moyenne (et erreur standard) de chaque variable mesurée en fonction du genre des élèves, du genre de l'agent, et du comportement de l'agent (sans ou avec rapport virtuel). Le Tableau 7.2 présente les corrélations bivariées entre

TABLEAU 7.1

Étude pilote 1A. Score moyen (erreur standard) pour l'ensemble des variables en fonction du genre de l'élève, du genre de l'agent et du comportement de l'agent avec ou sans rapport virtuel.

	Garçons		Filles	
Variables	Agent masculin	Agent féminin	Agent masculin	Agent féminin
Agent *sans* rapport virtuel				
Compétence en mathématiques	4.64 (.18)	4.62 (.15)	4.79 (.16)	4.22 (.20)
Sentiment de rapport	3.60 (.18)	3.81 (.15)	3.48 (.16)	3.19 (.20)
Sympathie	4.15 (.22)	4.17 (.19)	4.27 (.20)	3.87 (.25)
Anthropomorphisme	3.14 (.22)	2.96 (.19)	2.83 (.20)	2.51 (.24)
Identification à l'agent	3.50 (.29)	3.23 (.25)	3.29 (.26)	3.08 (.32)
Identification à la réussite de l'agent	4.75 (.16)	4.80 (.13)	4.13 (.14)	4.50 (.17)
Agent *avec* rapport virtuel				
Compétence en mathématiques	4.58 (.18)	4.78 (.15)	4.60 (.16)	4.37 (.20)
Sentiment de rapport	3.99 (.17)	4.22 (.15)	3.88 (.15)	3.79 (.19)
Sympathie	4.60 (.17)	4.45 (.15)	4.54 (.15)	4.56 (.19)
Anthropomorphisme	3.25 (.22)	2.91 (.19)	3.29 (.20)	2.65 (.25)
Identification à l'agent	3.71 (.25)	3.35 (.21)	3.46 (.22)	3.25 (.27)
Identification à la réussite de l'agent	4.59 (.18)	4.87 (.16)	4.00 (.16)	4.47 (.20)

l'ensemble des variables.

Dans toutes les conditions, les moyennes de la compétence des agents en mathématiques étaient supérieures à 4 (sur une échelle en 5 points), ce qui indique que tous les élèves ont perçu les agents comme très compétents en mathématiques. Ce constat est renforcé par le résultat au test de Student de comparaison au point médian (3) de l'échelle de Likert, M=4.60, ES=.08, $t(48)$=21.20, p<.001. De même, les moyennes de sympathie perçue et d'identification à la réussite en mathématiques des agents étaient supérieures ou égales à 4 dans toutes les conditions. Le même test de Student a montré que les élèves ont trouvé les agents très sympathiques, M=4.33, ES=.08, $t(48)$=17.15, p<.001, et ils se sont tous fortement identifiés à la réussite des agents M=4.51, ES=.08, $t(48)$=17.89, p<.001. Bien que de manière plus modérée, les élèves se sont également identifiés aux agents eux-mêmes, M=3.36, ES=.11, $t(48)$=3.17, p=.003, et ont perçu le rapport virtuel avec les agents, M=3.77, ES=.07, $t(48)$=10.67, p<.001. Le niveau d'anthropomorphisme est la seule variable dont la moyenne ne différait pas significativement du point médian, M=2.96, ES=.10, $t(48)$=-.37, p=711, indiquant que les élèves ont bien considéré que l'agent était proche de l'humain, mais pas trop.

Nous avons réalisé des analyses de corrélations bivariées entre les différentes mesures, et ce pour chacun des deux types de comportements des agents (sans *versus* avec rapport) afin de comparer la taille des corrélations entre les deux conditions (Tableau 7.2). Les analyses ont révélé une corrélation négative significative entre l'ordre

TABLEAU 7.2

Étude pilote 1A. Corrélations bivariées entre l'ensemble des variables dans la condition sans rapport virtuel (panel A) et dans la condition avec rapport virtuel (panel B).

Panel A.

Variables	M	ES	1	2	3	4	5	6	7	8	9
1. Ordre	-	-	-								
2. Sexe de l'élève	-	-	-.02	-							
3. Genre de l'agent	-	-	.02	-.19	-						
4. Compétence en mathématiques	4.60	.09	.04	-.05	-.21	-					
5. Sentiment de rapport	3.55	.09	-.13	-.29*	.04	.45**	-				
6. Sympathie	4.14	.10	.08	-.03	-.11	.40**	.82**	-			
7. Anthropo-morphisme	2.88	.11	-.06	-.23	-.12	.30*	.49**	.49**	-		
8. Identification à l'agent	3.28	.14	-.31*	-.07	-.11	.26	.58**	.63**	.53**	-	
9. Identification à la réussite de l'agent	4.54	.08	-.11	-.44**	.25	.32*	.42**	.28	.34*	.38**	-

Panel B.

Variables	M	ES	1	2	3	4	5	6	7	8	9
1. Ordre	-	-	-								
2. Sexe de l'élève	-	-	-.02	-							
3. Genre de l'agent	-	-	.02	-.19	-						
4. Compétence en mathématiques	4.61	.08	-.04	-.16	.03	-					
5. Sentiment de rapport	3.99	.08	.13	-.24	.12	.64**	-				
6. Sympathie	4.53	.08	-.03	.03	-.07	.46**	.56**	-			
7. Anthropo-morphisme	3.05	.11	.19	-.01	-.30*	.36**	.51**	.54**	-		
8. Identification à l'agent	3.44	.11	.02	-.08	-.16	.53**	.75**	.60**	.64**	-	
9. Identification à la réussite de l'agent	4.48	.10	-.07	-.42**	.34*	.57**	.53**	.42**	.14	.50**	-

Note. Les variables catégorielles ont été codées de la manière suivante : Ordre : 1=agent avec rapport suivi de l'agent sans rapport, 2=agent sans rapport suivi de l'agent avec rapport. Sexe de l'élève : 1=garçon, 2=fille. Genre de l'agent : 1=masculin, 2=féminin.

*p<.05. **p<.01.

de présentation de l'agent et le niveau d'identification à l'agent, et ce uniquement dans la condition où l'agent n'adopte pas de comportement pour instaurer un rapport virtuel (panel A). Le niveau d'identification à l'agent sans rapport virtuel était plus important lorsque cet agent a été vu en second, c'est-à-dire après l'avoir vu adopter des comportements favorisant le rapport virtuel. Cet effet d'ordre est finalement assez trivial : comme c'est le même agent dans les deux conditions, on ne peut oublier ses comportements préalables. Le fait que cette corrélation soit non significative dans la condition où l'agent adopte des comportements propices au rapport virtuel corrobore ce constat.

Toujours dans la condition de l'agent sans rapport virtuel (panel A), les garçons ont plus ressenti que les filles une certaine forme de rapport virtuel. Cette différence n'existait plus lorsque l'agent adoptait clairement des comportements pour favoriser le rapport virtuel (panel B). La seule différence liée au sexe de l'élève observable dans les deux conditions était une plus grande identification à la réussite en mathématiques de l'agent de la part des garçons que des filles. Concernant le genre de l'agent, il n'y avait aucune corrélation significative dans la condition où l'agent n'adoptait pas de comportements propices à l'instauration d'un rapport virtuel avec l'élève. En revanche, dans la condition avec comportements de rapport virtuel, bien que l'agent féminin était perçu comme moins proche d'un être humain que l'agent masculin, il semblait davantage faciliter chez les élèves l'identification à la réussite en mathématiques.

Bien que les corrélations entre le sentiment de rapport virtuel et les autres caractéristiques de l'agent (à l'exception de son genre) étaient toutes significatives dans les deux conditions (avec *versus* sans rapport), l'adoption de comportements de rapport virtuel changeait la taille de ces corrélations pour certaines variables. Les comportements de rapport virtuel ont renforcé la corrélation entre la perception du rapport virtuel et la perception du niveau de compétence en mathématiques, ainsi que l'identification à la fois à l'agent et à sa réussite en mathématiques. La corrélation entre le niveau de compétence en mathématiques de l'agent et le comportement de l'agent est passé de .45 à .64 lorsque l'agent démontrait des comportements de rapport virtuel. Autrement dit, le rapport virtuel aide à percevoir l'agent comme un modèle de réussite en mathématiques auquel on a envie de s'identifier. Dans le même temps, ça le rend un peu moins sympathique.

7.2.4.2. Analyses centrales

Il s'agit maintenant d'approfondir les observations précédentes en examinant si la perception différenciée de l'agent selon qu'il adopte ou non des comportements de rapport virtuel a été influencée par le sexe des élèves et/ou le genre de l'agent. L'analyse correspondante est une ANOVA mixte 2 (sexe des élèves) X 2 (genre de l'agent) X 2 (rapport virtuel), avec la dernière variable comme facteur intrasujets. Toutefois, compte tenu du faible nombre de participants dans l'une des conditions

TABLEAU 7.3

Étude pilote 1A. Score moyen et erreur standard pour chaque variable dépendante, en fonction du comportement de l'agent (avec ou sans rapport).

Variables	Agent sans rapport virtuel		Agent avec rapport virtuel		F	p	η^2_p
	M	ES	M	ES			
Compétence en mathématiques de l'agent	4.57	.09	4.58	.09	.03	.872	<.01
Sentiment de rapport	3.52	.09	3.70	.09	24.27	<.001	.35
Sympathie	4.11	.11	4.54	.08	16.40	<.001	.27
Anthropomorphisme	2.86	.11	3.02	.11	3.64	.063	.08
Identification à l'agent	3.28	.14	3.44	.12	2.40	.129	.05
Identification à la réussite en maths de l'agent	4.54	.08	4.48	.09	1.00	.324	.02

du plan expérimental (seulement 9 filles ont vu les agents féminins), les résultats sont à prendre avec précaution [1]. Les analyses n'ont révélé aucun effet significatif du sexe des élèves et du genre de l'agent, que ce soit en effet principal ou en interaction avec les autres variables. Seul un effet principal du comportement de l'agent a été obtenu pour le sentiment de rapport virtuel, la sympathie, et dans une moindre mesure (effet tendanciel) pour le niveau d'anthropomorphisme perçu (cf Tableau 7.3). Comme attendu, les élèves ont perçu plus de rapport virtuel de la part des agents avec rapport virtuel (M=3.70, ES=.09) en comparaison aux agents sans rapport virtuel (M=3.52, ES=.09), qui s'accompagnait par une plus grande perception de sympathie (avec rapport virtuel : M=4.54, ES=.08 ; sans rapport virtuel : M=4.11, ES=.11), et de manière tendancielle par un niveau d'anthropomorphisme plus élevé (avec rapport virtuel : M=3.02, ES=.11 ; sans rapport virtuel : M=2.86, ES=.11).

Enfin, des analyses de régressions multiples ont été réalisées séparément pour les conditions avec ou sans rapport virtuel, afin de mieux cerner les relations entre les variables liées à la perception de l'agent (cf. Tableau 7.4). Trois séries d'analyses ont été effectuées. Une première série avait pour objectif d'examiner si la perception du niveau de compétence en mathématiques de l'agent était influencée par le sentiment perçu de rapport virtuel, d'anthropomorphisme, et de sympathie. Les deux autres séries d'analyses avaient pour objectif de tester, respectivement, si l'identification à l'agent et l'identification à la réussite en mathématiques de l'agent, étaient influencées par l'ensemble des variables précédentes.

1. Il est à noter que des analyses alternatives ont été réalisées séparément pour le sexe de l'élève et pour le genre de l'agent (avec à chaque fois le comportement de l'agent en facteur intrasujets), afin de s'assurer que des effets liés au sexe de l'élève ou au genre de l'agent en interaction avec le comportement de l'agent n'aient pas été sous-estimés lors de l'ANOVA mixte 2X2X2 précédente. Ces analyses n'ont révélé aucun effet significatif autre que l'effet principal du comportement de l'agent déjà observé avec l'ANOVA 2x2X2.

TABLEAU 7.4

Étude pilote 1A. Analyses de régressions multiples réalisées séparément dans la condition sans rapport virtuel versus avec rapport virtuel.

Variables	Sans rapport virtuel	Avec rapport virtuel
	VD : Compétence en mathématiques de l'agent	
Sentiment de rapport virtuel	*B*=.35, *p*=.145	***B*=.56, *p*<.001**
Sympathie	*B*=.07, *p*=.779	*B*=.15, *p*=.316
Anthropomorphisme	*B*=.10, *p*=.541	*B*=-.01, *p*=.997
	VD : Identification à l'agent	
Sentiment de rapport virtuel	*B*=.13, *p*=.508	***B*=.51, *p*<.001**
Sympathie	***B*=.40, *p*=.045**	*B*=.14, *p*=.232
Anthropomorphisme	***B*=.28, *p*=.035**	***B*=.30, *p*=.007**
Compétence en mathématiques de l'agent	*B*=-.05, *p*=.708	*B*=.03, *p*=.801
	VD : Identification à la réussite de l'agent	
Sentiment de rapport virtuel	*B*=.45, *p*=.065	*B*=.13, *p*=.502
Sympathie	*B*=-.36, *p*=.137	*B*=.18, *p*=.235
Anthropomorphisme	*B*=.12, *p*=.461	***B*=-.37, *p*=.019**
Compétence en mathématiques de l'agent	*B*=.16, *p*=.272	***B*=.35, *p*=.022**
Identification à l'agent	*B*=.24, *p*=.185	*B*=.34, *p*=.086

Note. VD : variable dépendante.

Ces trois séries d'analyses de régressions permettent de compléter et de nuancer les résultats précédents. Premièrement, parmi les trois variables testées pour prédire la compétence perçue de l'agent en mathématiques (i.e., sentiment de rapport virtuel, sympathie, anthropomorphisme), seul le sentiment de rapport virtuel s'est révélé être un prédicteur significatif, et uniquement lorsque l'agent adoptait des comportements de rapport virtuel. Autrement dit, c'est la perception du rapport virtuel instauré par l'agent qui prime pour percevoir l'agent comme compétent en mathématiques. Deuxièmement, s'agissant de la prédiction de l'identification à l'agent, le niveau d'anthropomorphisme perçu apparaît comme un prédicteur important, que l'agent adoptait ou non des comportements propices au rapport virtuel. Par ailleurs, en l'absence de comportements de rapport virtuel, le niveau de sympathie perçu de l'agent a prédit l'identification de l'élève à l'agent alors qu'en présence de comportements de rapport virtuel, cette variable est devenue non pertinente et c'était la perception de rapport virtuel qui prédisait l'identification à l'agent. Ceci est également cohérent avec les résultats précédents qui nous ont indiqué que les comportements de rapport virtuel "rendaient" l'agent un peu moins sympathique. Enfin, il est intéressant de noter que la compétence perçue en mathématiques de l'agent n'a pas prédit l'identification de l'élève à l'agent, et ce que l'agent adoptait ou non des comportements de rapport virtuel.

Le troisième et dernier modèle testé concernait la prédiction de l'identification à la réussite en mathématiques de l'agent par l'ensemble des autres variables. Lorsque

l'agent n'adoptait pas de comportements propices à l'établissement d'un rapport virtuel avec l'élève, aucune des variables ne prédisait significativement l'identification de l'élève à la réussite de l'agent. En revanche, lorsque l'agent adoptait des comportements permettant d'instaurer un rapport virtuel, le niveau perçu de compétence en mathématiques et le degré d'anthropomorphisme de l'agent prédisaient significativement l'identification à la réussite en mathématiques de l'élève à l'agent. Comme l'on pouvait s'y attendre, plus l'agent était perçu comme compétent en mathématiques et plus l'élève s'identifiait à sa réussite. Les résultats obtenus avec l'anthropomorphisme sont plus inattendus : moins l'élève trouvait l'agent humain, et plus il s'identifiait à sa réussite en mathématiques. Nous proposons deux interprétations à ce résultat qui semble contre-intuitif. Notre première interprétation réside dans le sentiment d'anonymat conféré par un agent virtuel lorsqu'il exprime des comportements de rapport virtuel. Comme montré par DeVault et al. (2014), les agents virtuels, en tant que programmes informatiques à priori non contrôlés par un humain, offrent aux utilisateurs un certain degré d'anonymat qui les aide à se sentir plus à l'aise et à divulguer des informations personnelles sans craindre d'être jugés, en particulier lorsque ces agents ont des comportements de rapport virtuel. Le rapport virtuel pousse également les individus à plus se confier. Ceci pourrait pousser les élèves à ne pas avoir peur d'être jugés s'ils affirment ressentir l'envie de réussir en mathématiques comme l'agent virtuel, même s'ils n'ont pas de bonnes notes. Cela expliquerait pourquoi on ne retrouve pas ce résultat pour les agents virtuels sans comportements de rapport. Cependant, ceci implique qu'il y aurait un lien entre le niveau d'anthropomorphisme perçu et le fait de comprendre que l'agent virtuel n'est pas contrôlé par un humain. Notre deuxième interprétation est liée aux observations faites avec les corrélations bivariées. Dans la condition avec comportements de rapport virtuel, l'agent féminin était perçu comme moins proche d'un être humain que l'agent masculin, mais cet agent féminin facilitait davantage chez les élèves l'identification à la réussite en mathématiques. Le fait de plus s'identifier à la réussite d'un agent avec comportements de rapport virtuel lorsqu'on le trouve moins humain pourrait donc être dû à l'agent féminin. Enfin, il est intéressant de souligner que ce dernier modèle montre que le degré d'identification de l'élève à l'agent ne prédit pas le degré d'identification de l'élève à la réussite en mathématiques de l'agent lorsque les autres variables sont statistiquement contrôlées. Autrement dit, l'identification à l'agent en tant que personnage n'implique pas l'identification à sa réussite en mathématiques.

En résumé, l'ensemble de ces résultats indique que le rapport virtuel aide les élèves à percevoir l'agent comme un modèle de réussite en mathématiques auquel ils peuvent plus facilement s'identifier.

7.2.5. Discussion

Les élèves ont bien trouvé que tous les agents virtuels étaient compétents en mathématiques. De plus, ils s'identifiaient aux agents ainsi qu'à leur réussite en mathématiques, tout en les trouvant sympathiques. Les élèves ont trouvé les agents modé-

rément humains, comprenant qu'il s'agit d'un personnage virtuel. Aussi, les élèves ont bien perçu le rapport virtuel et l'ont plus ressenti lorsque l'agent exprimait des comportements de rapport virtuel. Plus précisément, l'adoption de comportements de rapport virtuel renforce les corrélations entre la perception du rapport virtuel et la perception du niveau de compétence en mathématiques, ainsi que l'identification à la fois à l'agent et à sa réussite en mathématiques. Les comportements de rapport virtuel permettent donc aux élèves de mieux percevoir l'agent comme un modèle de réussite en mathématiques, et de mieux s'y identifier. Il n'y a, de plus, aucune différence selon le genre des agents et le sexe des élèves. Une corrélation significative existe cependant entre le genre de l'agent et le niveau d'anthropomorphisme, ainsi qu'avec le niveau d'identification, uniquement dans la condition où l'agent exprime des comportements de rapport virtuel : l'agent masculin a été perçu comme plus proche de l'humain que l'agent féminin. L'agent féminin a quant à lui permis aux élèves de mieux s'identifier. Les résultats des analyses de régressions multiples montrant que, lorsqu'un agent démontre des comportements de rapport virtuel, moins l'on trouve cet agent humain et plus on s'identifie à sa réussite, pourrait donc être dû à l'agent féminin.

D'après nos analyses, le rapport virtuel permet de faire de nos agents de meilleurs modèles de réussite en mathématiques, ce qui est cohérent avec les études en sciences humaines où l'interaction et des liens plus personnels entre des femmes compétentes en sciences et des jeunes filles ont permis à ces dernières de mieux percevoir ces femmes scientifiques comme des modèles de réussite (BUCK et al., 2008). Cette étude pilote a été menée avec un plan mixte où les enfants ont vu le même agent deux fois, mais se comportant différemment (avec et sans rapport virtuel). Par conséquent, ils avaient un point de comparaison lorsqu'ils évaluaient ces agents. Aussi, ces évaluations reposaient uniquement sur la façon dont l'agent se présente pour permettre ce plan mixte. Il était en effet impossible de demander à un élève d'évaluer deux fois la totalité de l'interaction avec nos agents virtuels, comprenant sa présentation et l'explication d'une leçon de mathématiques. Cela aurait entraîné un sentiment de lassitude lié à la répétition et à la longueur de la scène. La prochaine étape consiste donc à évaluer nos agents dans un plan intersujets, comprenant cette fois-ci l'explication de la leçon, en plus de la présentation de l'agent. Cette configuration nous rapproche de notre étude principale, et nous permet d'explorer si les résultats de l'étude pilote 1A pourraient être reproduits dans cette approche intersujets.

7.3. Étude pilote 1B : perception des caractéristiques et capacités des agents virtuels pédagogiques à partir de la façon dont ils se présentent et de leurs explications de la leçon de mathématiques

7.3.1. Participants

Huit classes de REP+ (4 classes de 6e et 4 classes de 5e) ont participé à cette étude pilote, soit un total de N=95 élèves (51 garçons, 44 filles, M_{age}=11.73, ES=.09). Les participants n'étaient pas les mêmes élèves que pour l'étude pilote 1A présentée dans la section précédente.

7.3.2. Procédure

La procédure était la même que pour l'étude pilote 1A, à la différence qu'il a été précisé que les élèves évalueraient un personnage en 3D, et que ce personnage expliquerait une leçon de mathématiques. Nous leur avons précisé que, bien que le personnage en 3D parlerait d'un contrôle de mathématiques à faire, ils n'allaient pas passer de test de mathématiques.

7.3.3. Mesures

Les mêmes mesures que l'étude pilote précédente ont été utilisées : la **compétence en mathématiques** (α=.69), le **sentiment de rapport virtuel** (α=.86), le **niveau d'anthropomorphisme** (α=.75), la **sympathie** (α=.82), et l'**identification à l'agent** (α=.81) et l'**identification à la réussite en mathématiques de l'agent** (α=.67). Deux classes (28 élèves) n'ont pas complété l'échelle sur l'identification à la réussite en mathématiques de l'agent à cause d'une perte de wifi au moment des passations. Nous avons décidé de ne pas analyser ces résultats lors des ANOVA (en raison d'une insuffisance d'effectifs sur cette variable), mais de les conserver lors des corrélations et des analyses de régressions.

7.3.4. Résultats

7.3.4.1. Analyses descriptives

Le Tableau 7.5 présente la moyenne (et erreur standard) de chaque variable mesurée en fonction du genre des élèves, du genre de l'agent, et du comportement de l'agent (sans ou avec rapport virtuel). Le Tableau 7.6 présente les corrélations bivariées entre l'ensemble des variables.

TABLEAU 7.5

Étude pilote 1B. Score moyen (erreur standard) pour l'ensemble des variables en fonction du genre de l'élève, du genre de l'agent et du comportement de l'agent avec ou sans rapport virtuel.

	Garçons		Filles	
Variables	Agent masculin	Agent féminin	Agent masculin	Agent féminin
Agent *sans* rapport virtuel				
Compétence en mathématiques	4.41 (.14)	4.73 (.21)	4.37 (.22)	4.49 (.20)
Sentiment de rapport	3.43 (.14)	3.39 (.21)	3.03 (.22)	3.61 (.20)
Sympathie	4.26 (.15)	4.22 (.22)	3.73 (.23)	4.33 (.21)
Anthropomorphisme	2.91 (.15)	2.50 (.22)	2.65 (.23)	2.66 (.21)
Identification à l'agent	3.17 (.21)	3.43 (.31)	2.47 (.33)	3.36 (.30)
Identification à la réussite de l'agent	4.30 (.17)	4.28 (.20)	3.75 (.24)	4.36 (.19)
Agent *avec* rapport virtuel				
Compétence en mathématiques	4.33 (.22)	4.30 (.20)	4.37 (.21)	4.43 (.17)
Sentiment de rapport	3.65 (.22)	3.85 (.20)	3.31 (.21)	3.88 (.17)
Sympathie	4.00 (.23)	4.46 (.21)	4.18 (.22)	4.50 (.19)
Anthropomorphisme	2.54 (.23)	3.34 (.21)	2.21 (.22)	2.97 (.18)
Identification à l'agent	3.22 (.33)	3.46 (.30)	2.78 (.31)	3.45 (.26)
Identification à la réussite de l'agent	4.60 (.28)	4.53 (.22)	4.50 (.44)	4.65 (.20)

Dans toutes les conditions, les moyennes de la compétence des agents en mathématiques étaient supérieures à 4 (sur une échelle en 5 points), ce qui indique que tous les élèves ont perçu les agents comme très compétents en mathématiques. Ce constat a été renforcé par le résultat au test de Student de comparaison au point médian de l'échelle de Likert, M=4.43, ES=.06, $t(94)$=21.98, p<.001. Le même test de Student a montré que les élèves ont trouvé les agents très sympathiques, M=4.24, ES=.07, $t(94)$=17.26, p<.001, et ils se sont fortement identifiés à la réussite des agents, M=4.36, ES=.08, $t(66)$=17.25, p<.001. Ils ont eu tendance à s'identifier aux agents eux-mêmes, M=3.19, ES=.10, $t(94)$=1.86, p=.066. De manière modérée, les élèves ont ressenti le rapport virtuel avec les agents M=3.53, ES=.07, $t(94)$=7.62, p<.001. Le niveau d'anthropomorphisme est la seule variable dont la moyenne était en dessous du point médian et en différait significativement, M=2.76, ES=.08, $t(94)$=-3.15, p=002, indiquant que les élèves n'ont pas considéré les agents comme assez proches de l'humain. D'autres analyses (ci-dessous et en Section 7.3.4.2), notamment en prenant en compte le genre de l'agent ainsi que son comportement, et le sexe de l'élève, ont permis d'explorer si certains agents étaient perçus comme plus proches de l'humain selon ces conditions.

Nous avons effectué des corrélations bivariées pour chacun des comportements des agents (sans *versus* avec rapport) pour comparer la taille des corrélations entre les deux conditions (Tableau 7.6). Dans les deux conditions, le sexe de l'élève ne corrélait significativement avec aucune variable. En revanche, uniquement lorsque l'agent

TABLEAU 7.6

Étude pilote 1B. Corrélations bivariées entre l'ensemble des variables dans la condition sans rapport virtuel (panel A) et dans la condition avec rapport virtuel (panel B).

Panel A.

Variables	M	ES	1	2	3	4	5	6	7	8
1. Sexe de l'élève	-	-	-							
2. Genre de l'agent	-	-	.23	-						
3. Compétence en mathématiques	4.53	.09	-.07	.17	-					
4. Sentiment de rapport	3.45	.08	-.06	.16	.30*	-				
5. Sympathie	4.27	.10	-.13	.12	.19	.58**	-			
6. Anthropo-morphisme	2.71	.09	-.10	-.20	-.06	.33*	.21	-		
7. Identification à l'agent	3.21	.14	-.15	.23	.38**	.68*	.47**	.22	-	
8. Identification à la réussite de l'agent	4.22	.11	-.12	.15	.24	.54**	.51**	.06	.43**	-

Panel B.

Variables	M	ES	1	2	3	4	5	6	7	8
1. Sexe de l'élève	-	-	-							
2. Genre de l'agent	-	-	.03	-						
3. Compétence en mathématiques	4.73	.09	.06	.02	-					
4. Sentiment de rapport	4.11	.10	-.08	.28	.67**	-				
5. Sympathie	4.67	.06	.09	.29	.61**	.69**	-			
6. Anthropo-morphisme	3.20	.12	-.19	.45**	.41**	.64**	.69**	-		
7. Identification à l'agent	3.83	.13	-.09	.22	.61**	.70**	.70**	.72**	-	
8. Identification à la réussite de l'agent	4.59	.09	.07	.03	.57**	.65**	.54**	.16	.68**	-

Note. Les variables catégorielles ont été codées de la manière suivante : Sexe de l'élève : 1=garçon, 2=fille. Genre de l'agent : 1=masculin, 2=féminin.

*p<.05. **p<.01.

démontrait des comportements de rapport virtuel, le genre de l'agent corrélait avec le niveau d'anthropomorphisme : contrairement à l'étude pilote 1A, les élèves ont perçu l'agent féminin comme plus proche de l'humain que l'agent masculin.

Nous avons trouvé des résultats comparables à ceux de l'étude 1A : bien que les corrélations entre le sentiment de rapport virtuel et les autres caractéristiques de l'agent (à l'exception de son genre) étaient toutes significatives dans les deux conditions (avec *versus* sans rapport), l'adoption de comportements de rapport virtuel a augmenté la taille de ces corrélations. Les comportements de rapport virtuel ont renforcé la corrélation entre la perception du rapport virtuel et la perception du niveau de compétence en mathématiques, ainsi que la sympathie, le niveau d'anthropomorphisme, et l'identification à la fois à l'agent et à sa réussite en mathématiques. Ces comportements de rapport virtuel ont en effet permis de passer de .30 à .67 de corrélation entre les deux variables. Aussi, on remarque que les corrélations entre la compétence en mathématiques de l'agent et les autres variables (à l'exception du genre de l'agent) ont augmenté et sont devenues significatives lorsque l'agent montrait des comportements de rapport virtuel. En d'autres termes, le rapport virtuel facilite la perception de l'agent en tant que modèle de réussite en mathématiques, suscitant ainsi l'identification à sa réussite et à lui-même. Contrairement à l'étude 1A, cela l'a rendu plus sympathique. Les comportements de rapport virtuel étaient positivement corrélés au niveau d'anthropomorphisme des agents.

7.3.4.2. Analyses centrales

Nous souhaitons examiner si la perception des agents virtuels est influencée par leur genre, leur comportement, ou bien encore le sexe de l'élève. Tout comme dans l'étude 1A, nous avons effectué une ANOVA 2 (genre des élèves) X 2 (genre de l'agent) X 2 (rapport virtuel). Cependant, nous préconisons de tenir compte du faible nombre de participants dans une condition (9 filles ont vu l'agent masculin sans comportements de rapport virtuel) [2]. Un effet principal du genre de l'agent a été trouvé pour toutes les variables dépendantes, à l'exception de la compétence en mathématiques perçue de l'agent virtuel où aucun effet significatif (ni principal ni d'interaction) n'a été obtenu (cf Tableau 7.7). Ainsi, les élèves ont trouvé l'agent féminin plus sympathique et plus humain que l'agent masculin. De plus, ils ont ressenti plus de rapport virtuel avec un agent féminin, s'accompagnant par une plus grande identification [3]. Le comportement de l'agent a révélé un effet principal uniquement sur le degré de rapport virtuel ressenti, où les élèves ont bien ressenti plus de rapport virtuel lorsque l'agent démon-

2. Il est à noter que, tout comme pour l'étude pilote 1A, des analyses alternatives ont été réalisées séparément pour le sexe de l'élève et pour le genre de l'agent, avec à chaque fois le comportement de l'agent comme variable indépendante, afin de s'assurer que des effets liés au sexe de l'élève ou de l'agent en interaction avec le comportement de l'agent n'aient pas été sous-estimés lors de l'ANOVA mixte 2X2X2 précédente. Ces analyses n'ont révélé aucun effet significatif autre que les effets déjà observés avec l'ANOVA 2x2X2.

3. Il s'agit de l'identification à l'agent, et non à l'identification à la réussite en mathématiques de l'agent qui n'est pas analysée ici en raison du manque de participants.

TABLEAU 7.7

Étude pilote 1B. Score moyen et erreur standard pour chaque variable dépendante, en fonction du genre de l'agent (féminin ou masculin).

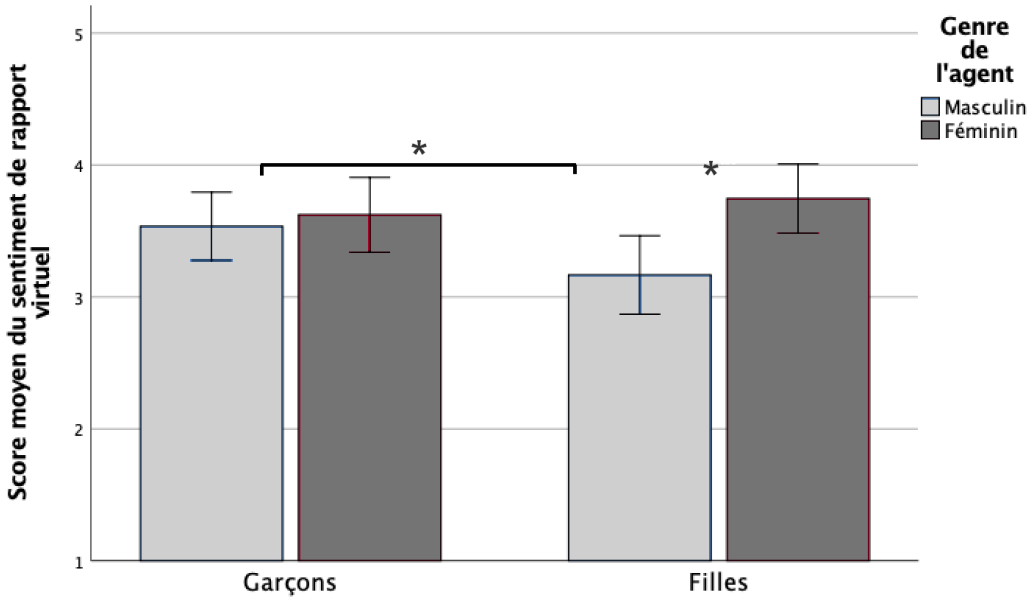
Variables	Agent masculin		Agent féminin		F	p	η_p^2
	M	ES	M	ES			
Compétence en mathématiques de l'agent	4.37	.10	4.49	.10	.72	.399	<.01
Sentiment de rapport	3.35	.10	3.68	.10	5.77	.018	.06
Sympathie	4.04	.11	4.38	.10	5.14	.026	.06
Anthropomorphisme	2.58	.11	2.87	.10	3.88	.052	.04
Identification à l'agent	2.91	.15	3.42	.15	6.05	.016	.07

trait des comportements de rapport virtuel (M=3.67, ES=.10 ; sans rapport : M=3.36, ES=.10), $F(1, 87)$=4.94, p=.029, η_p^2=.05.

Plus précisément, un effet d'interaction tendanciel entre le sexe de l'élève et le genre de l'agent a été obtenu sur le niveau de rapport virtuel ressenti, $F(1, 87)$=3.14, p=.080, η_p^2=.04. Les filles ont ressenti plus de rapport virtuel avec un agent féminin (M=3.75, ED=.13) comparé à un agent masculin (M=3.17, ED=.15), $F(1, 87)$=8.43, p=.005, η_p^2=.09. De manière cohérente, lorsque les élèves ont interagi avec un agent masculin, ce sont les garçons (M=3.54, ED=.13) qui ont ressenti plus de rapport virtuel que les filles (M=3.17, ED=.15), $F(1, 87)$=3.45, p=.066, η_p^2=.04. La Figure 7.1 illustre cette interaction.

FIGURE 7.1.

Étude pilote 1B. Score moyen du sentiment de rapport virtuel en fonction du sexe de l'élève (fille ou garçon) et du genre de l'agent (féminin ou masculin).

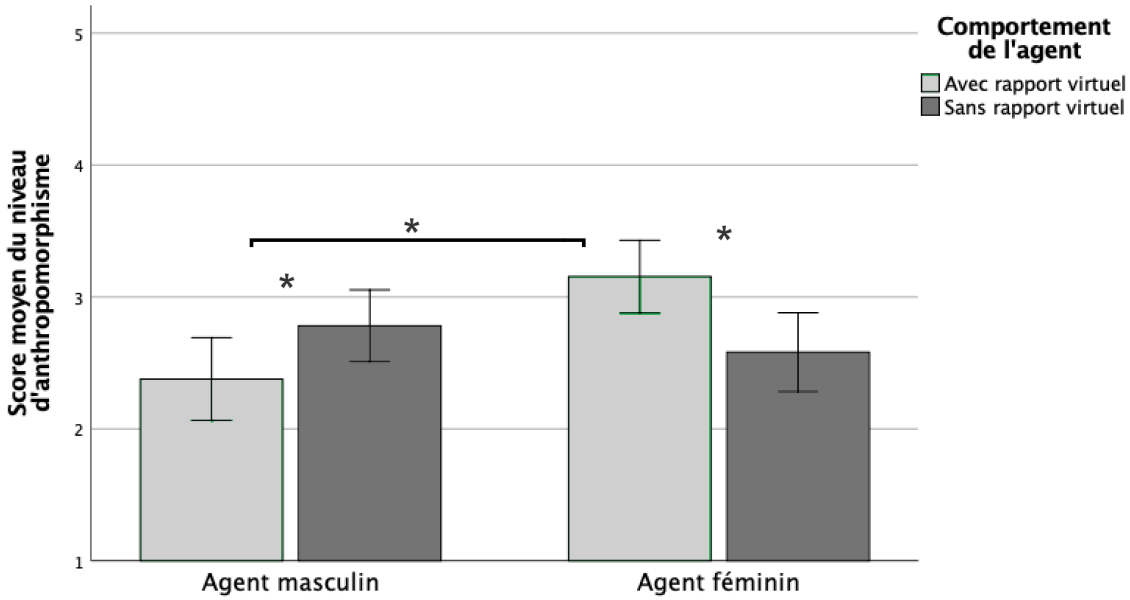


Note. Le symbole * représente les différences significatives. Barres d'erreur : intervalle de confiance à 95%.

Les analyses ont également révélé un effet d'interaction entre le genre de l'agent et son comportement sur le niveau d'anthropomorphisme, $F(1, 87)=11.13$, $p=.001$, $\eta^2_p=.14$. La décomposition de l'interaction avec correction de Bonferonni indique d'une part que, lorsque les agents démontraient des comportements de rapport virtuel, les agents masculins étaient perçus comme moins humains (rapport virtuel : $M=2.38$, $ES=.16$; sans rapport virtuel : $M=2.78$, $ES=.14$), $F(1, 87)=3.73$, $p=.057$, $\eta^2_p=.04$. C'est l'inverse pour les agents féminins, qui étaient perçus comme plus humains avec des comportements de rapport virtuel ($M=3.15$, $ES=.14$) plutôt que sans ($M=2.58$, $ES=.15$), $F(1, 87)=7.82$, $p=.006$, $\eta^2_p=.08$. Le niveau d'anthropomorphisme de l'agent féminin avec rapport virtuel ($M=3.15$, $ES=.14$) était plus élevé que pour l'agent masculin avec les mêmes comportements de rapport virtuel ($M=2.38$, $ES=.16$), $F(1, 87)=13.62$, $p<.001$, $\eta^2_p=.14$. Cette interaction est présentée en Figure 7.2.

FIGURE 7.2.

Étude pilote 1B. Score moyen du niveau d'anthropomorphisme en fonction du comportement de l'agent (avec ou sans rapport) et du genre de l'agent (féminin ou masculin).



Note. Le symbole * représente les différences significatives. Barres d'erreur : intervalle de confiance à 95%.

Enfin, comme pour l'étude pilote 1A, des analyses de régressions multiples ont été réalisées séparément pour les conditions avec ou sans rapport virtuel, afin de mieux cerner les relations entre les variables liées à la perception de l'agent (cf. 7.8). Trois séries d'analyses ont été effectuées. Une première série avait pour objectif d'examiner si la perception du niveau de compétence en mathématiques de l'agent était prédite par le sentiment perçu de rapport virtuel, la perception d'anthropomorphisme et de sympathie. Les deux autres séries d'analyses avaient pour objectif de tester, respectivement, si l'identification à l'agent et l'identification à la réussite en mathématiques de l'agent, étaient prédites par l'ensemble des variables précédentes. Nous avons retiré les 28 élèves n'ayant pas de données pour l'identification à la réussite en mathématiques de l'agent.

Ces trois séries d'analyses de régressions permettent de compléter et de nuancer les résultats précédents. Nous retrouvons les mêmes résultats que l'étude pilote 1A pour le premier modèle pour prédire la compétence perçue en mathématiques. Parmi les trois variables testées (i.e., sentiment de rapport virtuel, sympathie, anthropomorphisme), le sentiment de rapport virtuel était le seul prédicteur significatif, uniquement lorsque l'agent exprimait des comportements de rapport virtuel. Pour percevoir l'agent comme compétent en mathématiques, c'est la perception du rapport virtuel (instauré par l'agent avec des comportements de rapport virtuel) qui prime.

TABLEAU 7.8

Étude pilote 1B. Analyses de régressions multiples réalisées séparément dans la condition sans rapport virtuel versus avec rapport virtuel.

Variables	Sans rapport virtuel	Avec rapport virtuel
	VD : Compétence en mathématiques de l'agent	
Sentiment de rapport virtuel	*B*=.30, *p*=.117	***B*=.44, *p*=.049**
Sympathie	*B*=.05, *p*=.791	*B*=-.03, *p*=.902
Anthropomorphisme	*B*=-.18, *p*=.271	*B*=.03, *p*=.108
	VD : Identification à l'agent	
Sentiment de rapport virtuel	***B*=.49, *p*=.003**	*B*=.23, *p*=.246
Sympathie	*B*=.08, *p*=.614	***B*=.51, *p*=.020**
Anthropomorphisme	*B*=.05, *p*=.738	*B*=.09, *p*=.636
Compétence en mathématiques de l'agent	*B*=.21, *p*=.121	*B*=.07, *p*=.715
	VD : Identification à la réussite de l'agent	
Sentiment de rapport virtuel	***B*=.34, *p*=.068**	***B*=.28, *p*=.060**
Sympathie	***B*=.32, *p*=.040**	*B*=.26, *p*=.129
Anthropomorphisme	*B*=-.14, *p*=.312	*B*=-.25, *p*=.080
Compétence en mathématiques de l'agent	*B*=.05, *p*=.742	***B*=.34, *p*=.014**
Identification à l'agent	*B*=.12, *p*=.474	***B*=.41, *p*=.018**

Note. VD : variable dépendante.

Le deuxième modèle concerne la prédiction de l'identification à l'agent. Contrairement à l'étude pilote 1A, le niveau d'anthropomorphisme perçu n'est pas un prédicteur. Cette fois-ci, le niveau de sympathie perçue de l'agent prédit l'identification de l'élève à l'agent qui possède des comportements de rapport virtuel, alors qu'en absence de comportements de rapport virtuel, cette variable devient non pertinente et c'est la perception de rapport virtuel qui prédit l'identification à l'agent. Le seul résultat en commun avec l'étude pilote 1A est que la compétence perçue en mathématiques de l'agent ne prédit pas l'identification de l'élève à l'agent, que l'agent adopte des comportements de rapport ou non.

Le troisième et dernier modèle testé concerne la prédiction de l'identification à la réussite en mathématiques de l'agent par l'ensemble des autres variables. Lorsque l'agent n'adoptait pas de comportement propice à l'établissement d'un rapport virtuel avec l'élève, seul son niveau de sympathie perçu a prédit significativement l'identification de l'élève à la réussite de l'agent. Le sentiment de rapport virtuel perçu a tendance à prédire l'identification à la réussite, pour les agents avec ou sans comportements de rapport virtuel. Lorsque l'agent adoptait des comportements permettant d'instaurer un rapport virtuel, le niveau perçu de compétence en mathématiques et l'identification à l'agent prédisent significativement l'identification à la réussite en mathématiques de l'élève à l'agent. Plus l'agent était perçu comme compétent en mathématiques et plus l'élève s'identifiait à sa réussite. Contrairement à l'étude pilote

1A, ce dernier modèle montre que le degré d'identification de l'élève à l'agent prédit le degré d'identification de l'élève à la réussite en mathématiques de l'agent lorsque les autres variables sont statistiquement contrôlées. En d'autres termes, dans le cas où l'agent exprimait des comportements de rapport, le rapport virtuel a permis aux élèves de mieux comprendre que l'agent était compétent en mathématiques, ce qui leur a permis de mieux s'identifier à sa réussite en mathématiques.

7.3.5. Discussion

Dans cette étude pilote, l'agent s'est présenté et a rappelé les notions d'une leçon de mathématiques. Les élèves ont interagi avec un agent, contrairement à l'étude pilote précédente où les élèves ont interagi avec le même agent deux fois (le comportement variait), sans explications de la leçon de mathématiques. Dans cette étude pilote, le rapport virtuel a bien été perçu par les élèves et les agents ont été perçus comme sympathiques. Bien que tendancielle, les élèves arrivaient à distinguer un plus fort sentiment de rapport virtuel avec un agent qui adoptait des comportements propices à l'établissement d'un rapport virtuel. Il semblerait que le rapport virtuel soit plus dur à percevoir dans cette étude pilote par rapport à l'étude pilote précédente. Ceci peut être dû au point de comparaison présent dans l'étude pilote précédente, puisque le comportement de l'agent était une variable intrasujet. Une autre explication (qui n'exclue tout de même pas la première) serait que le sentiment de rapport virtuel s'estompait légèrement parce que l'agent donnait un cours de mathématiques. De ce fait, il s'intéressait un peu moins à l'élève en tant qu'individu comparé à la phase de présentation de l'agent, où il demandait par exemple comment l'élève s'en sortait en mathématiques. Ceci dit, dans la condition où l'agent adoptait des comportements de rapport virtuel, le sentiment de rapport virtuel ressenti a permis aux élèves de mieux percevoir les compétences en mathématiques de l'agent, ce qui a permis de mieux s'identifier à sa réussite.

Contrairement à l'étude pilote 1A, le rapport virtuel a été plus ressenti avec un agent féminin, qui a été également perçu comme plus sympathique qu'un agent masculin. Ceci peut être dû aux stéréotypes de genre : les agents féminins sont généralement notés comme plus sympathique que les agents masculins, mais cela n'explique pas la différence de résultat entre les deux étude pilotes. Il est également important de noter que ces stéréotypes pourraient être la cause des différences de niveau d'anthropomorphisme entre agents féminin et masculin, selon qu'ils véhiculent du rapport virtuel ou non. L'agent féminin avec rapport virtuel a été jugé comme plus humain que l'agent féminin sans rapport virtuel, alors que c'était l'inverse pour les agents masculins. Les agents avec rapport virtuel sourient plus et sont plus expressifs que les agents sans rapport. Il semblerait qu'un agent féminin doit être souriant et expressif, à l'inverse d'un agent masculin qui devrait être plutôt neutre, pour être considéré comme plus humain en reproduisant finalement les comportements stéréotypiquement genrés. De plus, dans la condition où l'agent exprime des comportements de rapport virtuel, un haut niveau d'anthropomorphisme a permis de s'identifier plus à l'agent. Ceci peut

être dû à l'agent féminin avec des comportements de rapport virtuel. Les élèves, filles comme garçons, se sont en effet plus identifiés à l'agent féminin qu'à l'agent masculin.

Contrairement aux stéréotypes de genre dans d'autres études perceptives où les participants (adultes) perçoivent les agents virtuels féminins moins experts que les agents virtuels masculins, les élèves ont perçu l'agent féminin et masculin comme étant tout aussi compétents en mathématiques l'un que l'autre. Nous pensons cependant qu'il serait hâtif de conclure que les enfants n'appliquent pas de stéréotypes de genre sur les mathématiques aux agents virtuels. Nous ne pouvons pas conclure que les enfants ne lient pas le genre des agents avec des compétences stéréotypées uniquement sur la base de nos résultats. Une étude a montré que les participants adultes n'attribuaient pas plus de traits de compétence aux agents masculins qu'aux agents féminins, cependant, ces agents étaient très similaires en apparence (NAG & YALÇIN, 2020). Puisque nous avons conçu l'agent masculin à partir de l'agent féminin et qu'ils portent les mêmes vêtements, il est possible que nos résultats soient du à leurs similitudes. Plus de recherche devrait être faite dans ce domaine, avec par exemple des agents très différents physiquement, avec différents degrés de féminité et masculinité. Au-delà du genre, bien qu'important, le niveau de masculinité perçu pourrait également jouer un rôle. Dans nos études pilotes, il n'était pas demandé aux participants d'évaluer le niveau de masculinité/féminité des agents, ce qui aurait pu être intéressant au vu des résultats. Enfin, le discours des agents sur leur réussite en mathématiques, couplé par l'explication d'une leçon de mathématiques, pourrait avoir montré aux élèves que les agents féminin et masculin sont tous les deux compétents en mathématiques. En d'autres termes, ces résultats pourraient être spécifiques à nos agents. Si tel est le cas, ce n'est pas un problème pour nous car nous les avons spécifiquement conçus afin de réduire les effets de menace du stéréotype.

7.4. Discussion

L'objectif commun aux deux études pilotes présentées dans ce chapitre était de vérifier que nos agents virtuels possédaient tous les prérequis pour incarner des modèles de réussite, afin de les utiliser dans l'étude principale dans le but de réduire les effets de menace du stéréotype. Il s'agissait de vérifier que tous les agents étaient bien perçus comme étant compétents en mathématiques et qu'il n'y avait pas de différence en fonction du genre de l'agent et du sexe des élèves. Également, nous voulions évaluer si le rapport virtuel était d'une part bien perçu par les élèves, et d'autre part si ce sentiment de rapport virtuel permettait de mieux évaluer le niveau de compétence de l'agent en mathématiques et de mieux s'y identifier. Les deux études pilotes se distinguaient par le traitement de la variable du comportement de l'agent (sans *versus* avec rapport virtuel). L'étude pilote 1A reposait sur un plan expérimental mixte 2 (genre de l'agent : féminin *versus* masculin) X 2 (comportement de l'agent : avec *versus* sans rapport virtuel), avec le genre en facteur intersujet et le rapport virtuel en facteur intrasujet. Les élèves évaluaient ainsi deux fois le même agent, mais avec des com-

portements différents. L'étude pilote 1B reposait quant à elle sur un plan intersujets 2 (genre de l'agent : féminin *versus* masculin) X 2 (comportement de l'agent : avec *versus* sans rapport virtuel). Pour éviter que les élèves se lassent en raison de la répétition et de la longueur de la scène, le rôle de l'agent dans l'étude 1A a été restreint à la seule séquence de sa présentation en tant que modèle de réussite en mathématiques. En revanche, l'étude 1B adoptant une approche intersujets, cela a permis d'inclure la séquence où l'agent expliquait la leçon. Cette approche aligne davantage l'étude 1B avec le protocole de l'étude principale. Ces études pilotes permettaient également d'explorer si l'ajout de la leçon renforçait ou diminuait les perceptions des élèves à l'égard des agents.

Les résultats des deux études pilotes sont encourageants car les élèves ont perçu tous les agents comme étant compétents en mathématiques, et se sont identifiés aux agents et à leur succès en mathématiques. Tous ces éléments sont des prérequis essentiels pour incarner un modèle de réussite, et ainsi avoir un impact positif sur les performances en mathématiques des filles, sans réduire celles des garçons. Ils ont également perçu le rapport virtuel, et ce sentiment de rapport virtuel que les élèves ont ressenti (uniquement lors de l'interaction avec les agents exprimant des comportements de rapport virtuel) a permis aux élèves de mieux percevoir les agents en tant que modèles de réussite. Plus précisément, les élèves les ont perçus comme plus compétents en mathématiques et se sont plus identifiés, à la fois à ces agents et à leur réussite en mathématiques. De plus, contrairement aux adultes dans la plupart des études dans la littérature sur les agents virtuels, les enfants n'ont pas appliqué de stéréotype de genre concernant la compétence en mathématiques perçue : l'agent masculin n'a pas été jugé comme plus compétent en mathématiques que l'agent féminin, dans les deux études pilotes. Des éléments de perceptions stéréotypiques ont tout de même été trouvés dans l'étude pilote 1B. Les élèves ont jugé l'agent féminin comme plus sympathique que l'agent masculin, et le niveau d'anthropomorphisme perçu des agents étaient plus ou moins haut selon le genre et le comportement de l'agent. L'agent féminin avec des comportements de rapport virtuel était jugé comme plus humain que lorsque cet agent n'avait pas de comportements de rapport virtuel. C'était l'inverse pour l'agent masculin. Également, les élèves ont ressenti plus de rapport virtuel avec l'agent féminin. Nos résultats apportent des informations supplémentaires aux recherches de BORAU et al. (2021), où les auteurs ont montré que les participants ont noté des images de robots féminins comme plus humaines que des images de robots masculins. Les comportements des agents et des robots couplés à leur apparence genrée peuvent modérer cette perception de l'humanité. Ceci étant dit, dans les deux études pilotes, les élèves ont bien ressenti plus de rapport virtuel avec l'agent exprimant des comportements de rapport virtuel (*versus* l'agent sans comportements de rapport virtuel), validant ainsi notre modèle de rapport virtuel présenté au Chapitre 6.

Une limite importante est cependant à prendre en compte. Le fait d'attribuer une condition à une classe a créé un déséquilibre du nombre de filles et de garçons, dû à

la parité des classes qui n'est pas toujours respectée. Ces études pilotes se déroulant directement au collège, où une classe entière évaluait les agents en même temps dans la même salle, il nous semblait toutefois plus judicieux d'attribuer une condition à une classe pour que les élèves voient tous les mêmes agents, et pour que la durée de la tâche soit la même pour une même classe. Si un élève avait vu un agent différent sur la tablette d'un camarade, cela aurait entraîné des interrogations de sa part, ce qui aurait pu perturber l'étude ou créer des biais dans la notation de cet élève. Nos résultats concernant les analyses de variances sont à prendre avec précaution, à cause de ce déséquilibre entre filles et garçons dans les conditions expérimentales.

Le résultat principal de ces deux études pilotes est que le rapport virtuel ressenti lors de l'interaction avec les agents exprimant des comportements de rapport virtuel aide à percevoir l'agent comme un modèle de réussite plus compétent en mathématiques auquel on a envie de s'identifier, à la fois à lui-même et à sa réussite. Autrement dit, le rapport virtuel permet de faire de nos agents de meilleurs modèles de réussite en mathématiques. Ces deux études pilotes valident la construction de nos agents, mais ont par elles-mêmes une contribution intéressante à la littérature scientifique dans le domaine, qui manque d'agents virtuels pour des enfants/pré-adolescents. À ce jour et à notre connaissance, c'est la première fois que de tels agents sont développés et dont on évalue auprès d'élèves la perception sur tout un ensemble de caractéristiques clés pour établir la possibilité d'une identification entre l'élève et l'agent. C'est également la première fois que le rapport virtuel dans une interaction agent-humain a été évalué et ressenti par des enfants/pré-adolescents. De plus, dans l'étude pilote 1B, les élèves se sont plus identifiés à l'agent féminin qu'à l'agent masculin. Des études ont montré que, dans le monde réel, les élèves ne s'identifiaient pas aux femmes dans un rôle de modèle de réussite lorsqu'elles étaient anxieuses quant à leurs capacités en mathématiques. Compte tenu de nos résultats sur l'identification après que l'agent ait expliqué la leçon de mathématiques (étude pilote 1B), on pourrait penser qu'un plus grand nombre d'enseignantes qui aiment les mathématiques, le font savoir, et démontrent leur compétence dans ce domaine produirait une plus grande identification à ces modèles de réussite. Nos résultats ouvrent la possibilité qu'un modèle de réussite virtuel féminin avec rapport virtuel soit non seulement efficace pour améliorer les performances en mathématiques des élèves, mais aussi pour encourager plus de filles dans les domaines des sciences en ayant une plus grande représentation.

7.5. Points clés du chapitre

- Validation des agents virtuels perçus comme modèle de réussite.
- Le rapport virtuel exprimé par nos agents pédagogiques a été évalué et perçu par des enfants/pré-adolescents.
- Le rapport virtuel ressenti lors de l'interaction avec les agents exprimant des comportements de rapport virtuel a aidé à percevoir ces agents comme des modèles de réussite plus compétents en mathématiques auxquels on a envie de s'identifier, à la fois à eux-mêmes et à leur réussite.
- Le rapport virtuel exprimé par nos agents pédagogiques a permis de faire d'eux de meilleurs modèles de réussite en mathématiques.
- Les élèves ont perçu les agents féminins et masculins tout aussi compétents en mathématiques l'un que l'autre.
- Lors d'une interaction plus longue avec l'agent, les filles ont ressenti plus de rapport virtuel avec l'agent féminin.
- Lors d'une interaction plus longue avec l'agent, l'agent féminin a été jugé comme plus humain lorsqu'il exprimait des comportements de rapport virtuel, et inversement pour l'agent masculin.

8. Étude pilote 2 : élaboration et évaluation des tests standardisés de mathématiques

Sommaire

8.1. Objectifs . . . 201
8.2. Matériels créés . . . 202
8.2.1. Classe de 6e . . . 202
8.2.2. Classe de 5e . . . 202
8.3. Participants . . . 203
8.4. Procédure . . . 203
8.5. Mesures et stratégies d'analyses descriptives . . . 204
8.6. Résultats . . . 205
8.6.1. Analyses descriptives globales . . . 205
8.6.2. Analyses descriptives chez les élèves de 6e . . . 206
8.6.2.1. Test 1 . . . 206
8.6.2.2. Test 2 . . . 207
8.6.3. Analyses descriptives chez les élèves de 5e . . . 210
8.6.3.1. Test 1 . . . 210
8.6.3.2. Test 2 . . . 212
8.7. Discussion . . . 214
8.8. Points clés du chapitre . . . 215

8.1. Objectifs

Ce chapitre détaille la manière dont nous avons élaboré les tests standardisés de mathématiques, en vue de les utiliser lors de l'étude principale pour mesurer l'impact de nos agents virtuels pédagogiques sur les effets de menace du stéréotype. La menace du stéréotype s'observe habituellement sur des exercices suffisamment difficiles et standardisés. Par conséquent, nous ne pouvons pas nous appuyer sur les pratiques individuelles des enseignants dans leurs classes : toutes les classes participantes doivent passer les mêmes tests de mathématiques. Nous avons collaboré avec cinq enseignants de mathématiques d'un collège en REP+, ainsi qu'avec un chargé d'éducation en zone prioritaire (également enseignant de mathématiques au collège) et un inspecteur académique en mathématiques pour concevoir les exercices et les leçons associées. Dans le cadre de notre étude principale, le rôle principal de l'agent virtuel pédagogique était de rappeler les notions clés d'une leçon de mathématiques. Ceci permettait de proposer la même expérience à tous les élèves, afin de ne pas dépendre uniquement de leurs connaissances préalables ni de la méthodologie de leur enseignant. Nous avons sélectionné les notions pour lesquelles les élèves rencontraient le plus de difficultés d'après l'équipe pédagogique, afin d'une part de créer des exercices difficiles, et d'autre part d'évaluer si nos agents virtuels pédagogiques peuvent aider les élèves là où ils en ont le plus besoin. L'équipe pédagogique avec laquelle nous avons travaillé nous a indiqué que les élèves de 6e rencontraient des difficultés avec les notions de parallélisme, tandis que ceux de 5e éprouvaient des difficultés avec les priorités opératoires. Les enseignants ont conçu les tests de mathématiques de manière à ce qu'ils soient assez difficiles, sans pour autant être irréalisables.

L'étude principale se déroulait en deux temps. Le premier temps servait à habituer les élèves à l'expérimentation numérique et à récolter quelques mesures en *baseline*. Ils n'interagissaient donc pas avec un agent, mais regardait à la place une vidéo explicative de la leçon de mathématiques. Les élèves interagissaient avec un agent lors du deuxième temps. Dans les deux cas, ils effectuaient un test de mathématiques. Nous avons donc créé, en étroite collaboration avec l'équipe pédagogique présentée ci-dessus, quatre tests de mathématiques : deux pour les 6e, et deux pour les 5e. Bien que les tests abordaient les mêmes notions dans les deux phases de l'expérimentation principale (le parallélisme pour les 6e et les priorités opératoires pour les 5e), ils étaient différents pour éviter les biais d'apprentissage. Nous avons prétesté ces tests de mathématiques afin d'ajuster le nombre d'exercices selon leur difficulté. Le but était d'avoir un équilibre entre les exercices faciles et difficiles, afin d'analyser l'impact des agents virtuels pédagogiques selon le niveau de difficulté des exercices. Il s'agit ici d'analyser de manière descriptive chaque question des quatre tests de mathématiques.

8.2. Matériels créés

8.2.1. Classe de 6e

La leçon de mathématiques à destination des 6e portait sur le parallélisme. La leçon présentait le vocabulaire du parallélisme, puis les propriétés. Dans un premier temps, la leçon a introduit les définitions des droites sécantes, du point d'intersection, des droites parallèles, et des droites perpendiculaires. Dans un second temps, trois propriétés en rapport avec les droites ont été expliquées. La propriété 1 permet de montrer que deux droites sont parallèles : si deux droites sont perpendiculaires à une même troisième droite, alors ces deux droites sont parallèles entre elles. La propriété 2 permet également de montrer que deux droites sont parallèles : si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles. Enfin, la propriété 3 permet de montrer que deux droites sont perpendiculaires : si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre. Des exemples ont été donnés à chaque étape. Enfin, la leçon se terminait sur un point méthode en donnant des conseils pour réussir son contrôle : (1) bien lire l'énoncé pour bien identifier les droites auxquelles on s'intéresse, (2) utiliser l'outil stylo de l'application pour repasser les droites avec différentes couleurs afin de mieux les identifier, (3) identifier à quelle figure-clé, parmi les trois figures se référant aux exemples des propriétés, correspond la figure de l'exercice, (4) utiliser la propriété correspondante à la figure-clé pour conclure. La vidéo de la leçon dure environ 3 minutes et 23 secondes.

Les deux tests de mathématiques comportaient des exercices en lien avec les notions du parallélisme. Le premier test contenait 7 exercices, et le deuxième en comptait 12. Il était demandé d'identifier la relation entre deux droites, d'utiliser les bons symboles pour exprimer cette relation, et d'utiliser une propriété du cours pour justifier la réponse. Les exercices prétestés sont donnés en Annexes F et G. Chaque figure représentant des droites ou des formes géométriques était mise à côté de l'exercice associé. Nous avons suivi les conseils des enseignants ayant créé les leçons et les tests, en créant un outil stylo avec différentes couleurs pour que les élèves puissent repasser sur les figures afin de mieux identifier les différentes droites.

8.2.2. Classe de 5e

Les 5e ont revu les notions de priorités opératoires. La leçon présentait le vocabulaire des opérations, puis comment calculer une expression, et enfin comment identifier une expression. Dans un premier temps, la leçon a introduit les définitions de la somme, de la différence, du produit, et du quotient. Dans un second temps, la leçon a expliqué par quelle opération commencer lorsqu'il n'y a pas de parenthèse, puis lorsqu'il y a des parenthèses. Dans un troisième temps, la façon d'identifier la nature d'une expression numérique (est-ce une somme, une différence, un produit, ou un quotient) a été montrée. Des exemples ont été donnés à chaque étape. Enfin,

la leçon se terminait sur un point méthode en donnant des conseils pour réussir son contrôle : (1) voir par quelle opération commencer pour calculer une expression, (2) ne pas oublier de calculer en priorité ce qu'il y a dans les parenthèses, en appliquant les règles des priorités opératoires (d'abord la multiplication ou la division, ensuite l'addition ou la soustraction, en calculant de gauche à droite), (3) voir la dernière opération à effectuer pour trouver la nature d'une expression. Cette vidéo dure environ 4 minutes et 19 secondes.

Les deux tests de mathématiques comportaient des exercices sur les priorités opératoires. Les élèves devaient calculer des expressions, identifier la nature des expressions, bien placer les signes d'opération, traduire une expression en français et inversement. Le premier test contenait 17 exercices, et le deuxième en comptait 16. Les exercices prétestés sont donnés en Annexes H et I. Les tests de 5e contenaient plus d'exercices que ceux de 6e car ils étaient adaptés à ce que les enseignants attendent des élèves selon leur classe.

8.3. Participants

Une classe de 6e (N=17, 8 filles, 9 garçons, M_{age}=11.00, ES=.09) et une classe de 5e (N=21, 10 filles, 11 garçons, M_{age}=12.14, ES=.08) en REP+ ont participé à cette étude pilote. Le collège participant à cette étude n'est pas le même collège où les enseignants ont créé les exercices de mathématiques.

8.4. Procédure

Les élèves se sont installés dans une salle informatique du collège avec leur enseignant de mathématiques, l'expérimentatrice, et une assistante d'éducation. Les ordinateurs étaient tous équipés de casque audio fourni par Ampiric (MIRAULT et al., 2023). La classe de 6e a passé l'expérimentation en premier, et la classe de 5e l'a passé en deuxième. En début de séance, l'expérimentatrice s'est présentée et a expliqué aux élèves la procédure à suivre. Ses explications étaient similaires à celles qu'elles donneraient lors de l'étude principale. Elle s'est présentée en tant que personne travaillant pour les collèges, et ayant été contactée par le chef d'établissement pour faire passer un bilan de compétences aux élèves. Les élèves ont été avertis que le test de mathématiques se passerait sur ordinateur, et compterait dans la moyenne. L'expérimentatrice a précisé qu'avec son équipe, elle avait conçu le site web où les élèves passeraient le test de mathématiques, mais que les tests de mathématiques avaient été conçus par les enseignants et les inspecteurs académiques. Les élèves étaient informés que l'évaluation en mathématiques comprenait deux parties. Dans un premier temps, ils ont visionné la vidéo de la leçon où une voix (féminine ou masculine) expliquait les notions clés d'une leçon de mathématiques. Ils effectuaient ensuite le test de mathématiques de 30 minutes sur les notions abordées dans la vidéo. Il était recommandé aux élèves d'être très attentifs à la leçon afin de bien réussir le test

de mathématiques. Les élèves pouvaient poser des questions sur le déroulement de la séance, ou s'ils n'avaient pas compris une tâche. Une fois les explications terminées, la classe a été répartie en demi-groupe dans les deux salles informatiques du collège. Un demi-groupe a été surveillé par l'enseignant, alors que l'autre demi-groupe a été surveillé par l'expérimentatrice et l'assistante d'éducation. Les élèves ont mis leur casque et ont commencé la séance. Ils pouvaient poser des questions en levant le doigt s'ils n'avaient pas compris un exercice ou s'ils rencontraient des difficultés liés au site web. Les élèves n'avaient pas le droit à la calculatrice, mais pouvait prendre une feuille de brouillon. Les élèves ayant terminé avant la fin de l'heure étaient invités à rester silencieux et à faire leurs devoirs, ou bien à continuer leur avancée de leurs cours de mathématiques sur le site de Studio Code (application en ligne pour apprendre les bases de la programmation, utilisée dans les classes). À la fin de chaque séance, les élèves étaient remerciés.

Les demi-groupes ont été composés de sorte à avoir autant de filles que de garçons, avec une moyenne en mathématiques comparable entre les groupes. Chaque groupe a de nouveau été coupé en deux, de manière à ce qu'il y ait autant d'élèves ayant visionné la vidéo de la leçon avec une voix féminine que d'élèves ayant visionné la même vidéo avec une voix masculine. Le Tableau 8.1 présente les effectifs avec leur moyenne trimestrielle en mathématiques.

TABLEAU 8.1

Nombre de participants avec leur moyenne trimestrielle en mathématiques (et erreur standard) par conditions expérimentales.

	Fille		Garçon	
	N	*M* (*ES*)	*N*	*M* (*ES*)
Test 1				
6e	4	13.86 (3.05)	4	13.63 (4.00)
5e	5	10.80 (3.91)	4	12.31 (2.41)
Test 2				
6e	4	14.05 (1.32)	5	12.08 (4.78)
5e	5	11.05 (4.06)	7	10.48 (5.47)

8.5. Mesures et stratégies d'analyses descriptives

Performance. Le score de performance de chaque question pour tous les tests de mathématiques a été analysé de manière descriptive, compte tenu du faible nombre de participants. Le score de chaque question est un pourcentage. Il s'agit ici d'identifier les exercices trop faciles ou trop difficiles, afin de trouver un équilibre entre le nombre d'exercices simples et compliqués. Nous avons également calculé la performance globale pour chaque test. Des tests de Student ont été réalisés pour examiner les

différences entre les filles et les garçons.

Temps de réponse global. Le temps de réponse global aux tests de mathématiques a été mesuré en minute afin de raccourcir le test s'il est trop long, ou de rajouter des exercices s'il est trop court.

Moyenne trimestrielle en mathématiques. Nous avons recueilli auprès de l'établissement scolaire la moyenne en mathématiques du premier trimestre de chaque élève, comprise entre 0 et 20. Ceci nous a permis de comparer les performances avec le niveau initial en mathématiques des élèves.

8.6. Résultats

8.6.1. Analyses descriptives globales

Le Tableau 8.2 présentent les moyennes trimestrielles en mathématiques des filles et des garçons, ainsi que leur performance (note convertie sur 20) aux tests standardisés de mathématiques.

TABLEAU 8.2

Moyennes (et erreurs standards) trimestrielle en mathématiques et obtenues lors des tests standardisés de mathématiques, pour les filles et les garçons.

	Filles (N=18)	Garçons (N=20)
M_{maths} (ES)	12.27 (.81)	11.88 (.98)
M_{tests} (ES)	10.22 (.63)	11.82 (.75)

Note. M_{maths} : moyenne trimestrielle en mathématiques. M_{tests} : moyenne des tests standardisés de mathématiques présentés dans ce chapitre.

Des tests de Student ont montré que la moyenne obtenue par les garçons lors des tests standardisés (M=11.82, ES=.75) était similaire à leur niveau initial en mathématiques (M=11.88, ES=.98), $F(19)$=.05, p=.962. En revanche, les filles ont moins bien réussi les tests (M=10.22, ES=.63) par rapport à leur moyenne trimestrielle en mathématiques (M=12.27, ES=.81), $F(17)$=2.31, p=.039. Il est fort probable que la menace du stéréotype ait joué un rôle dans ces résultats. Bien que les différences de résultats aux tests standardisés de mathématiques entre filles et garçons ne soient pas significatifs, $F(36)$=1.62, p=.115, les filles ont produit des contre-performances par rapport aux notes qu'elles obtiennent habituellement, contrairement aux garçons qui ont conservé les mêmes notes qu'en classe.

Les prochaines parties de cette section sont consacrées aux analyses descriptives de chaque exercice composant tous les tests standardisés de mathématiques. Ces analyses nous ont permis de prendre des décisions quant à l'ajustement des tests, en identifiant les exercices faciles et difficiles.

8.6.2. Analyses descriptives chez les élèves de 6e

8.6.2.1. Test 1

Le Tableau 8.3 présente la moyenne et l'erreur standard de chaque exercice des 6e pour le test 1. La Figure 8.1 montre les différences de performance entre les filles et les garçons pour ces mêmes exercices.

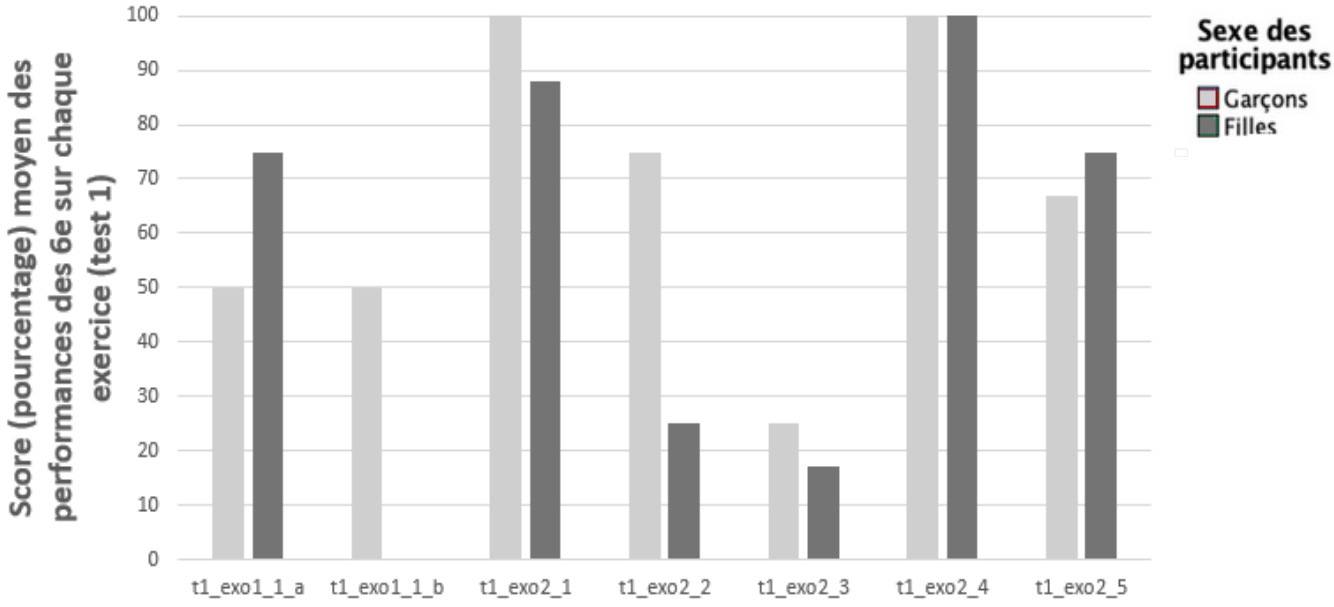
TABLEAU 8.3

Moyenne et erreur standard pour chaque exercice du test 1 des 6e.

Exercice	*M*	*ES*
t1_exo1_A	63.00	.18
t1_exo1_B	25.00	.16
t1_exo2_1	93.80	.06
t1_exo2_2	50.00	.19
t1_exo2_3	20.83	.06
t1_exo2_4	100.00	.00
t1_exo2_5	70.83	.15

FIGURE 8.1.

Score moyen des performances des garçons et des filles de 6e pour chaque exercice du test 1.

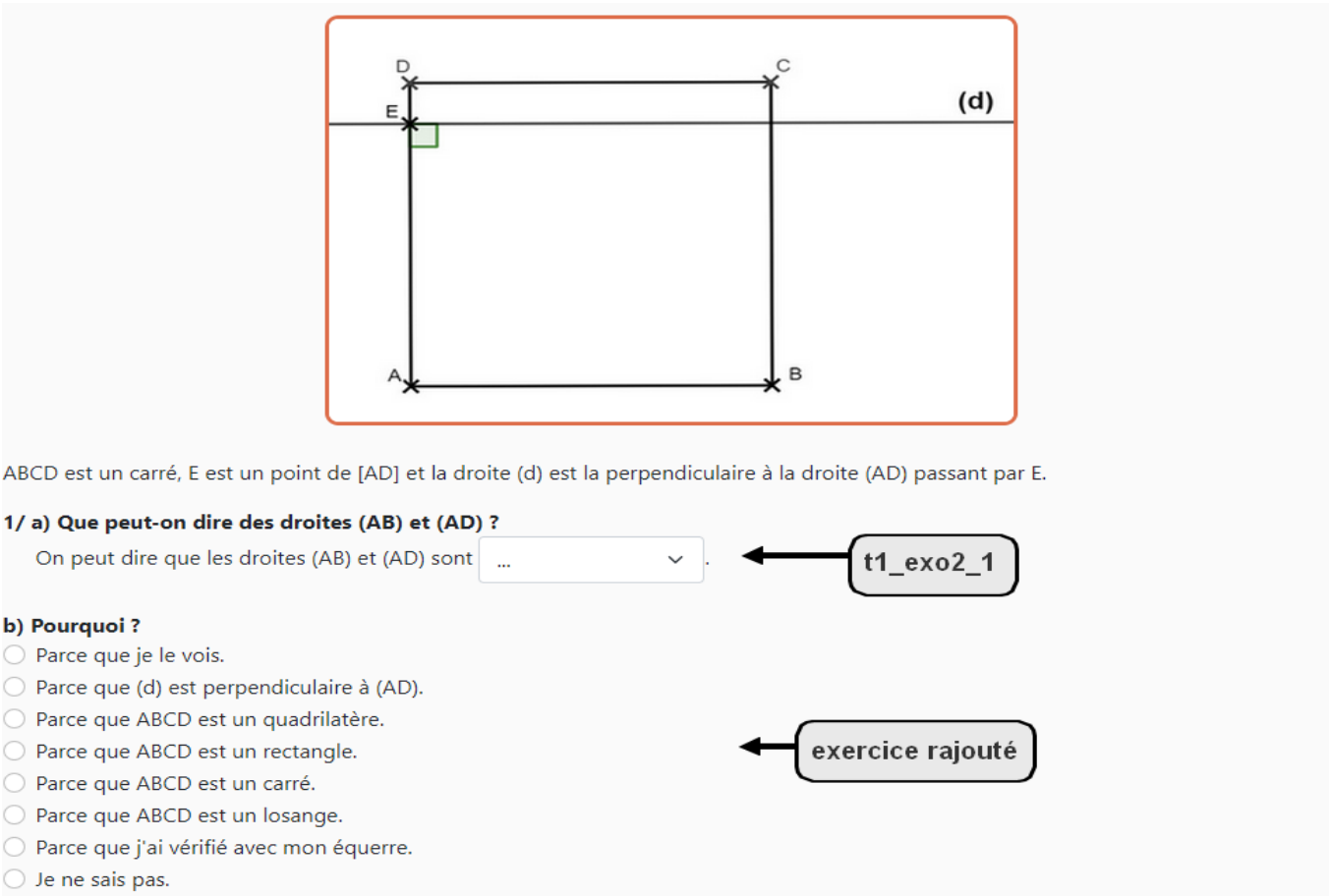


Bien que tout le monde ait réussi l'exercice *t1_exo2_4*, nous avons décidé de le garder d'une

part pour avoir un équilibre entre le nombre d'exercices trop simples et trop compliqués, et d'autre part pour ne pas raccourcir le temps de réponse global au test qui était assez bas (M=7.03, ES=.63). Pour les mêmes raisons, nous avons rajouté un exercice jugé difficile par les enseignants où il est demandé de justifier la réponse de l'exercice *t1_exo2_1*. L'exercice *t1_exo2_1* ainsi que l'exercice rajouté sont donnés en Figure 8.2. Globalement, les garçons ont obtenu en moyenne 56.67% de bonnes réponses (ES=.20), et les filles 47.50% (ES=.48). Les filles (M=6.15, ES=.73) comme les garçons (M=7.90, ES=2.22) ont fini le test très rapidement.

FIGURE 8.2.

Exercice (jugé difficile par les enseignants) rajouté au test 1 de 6e.



8.6.2.2. Test 2

Le Tableau 8.3 présente la moyenne et l'erreur standard de chaque exercice du test 2. La Figure 8.1 montre les différences de performance entre les filles et les garçons pour ces exercices.

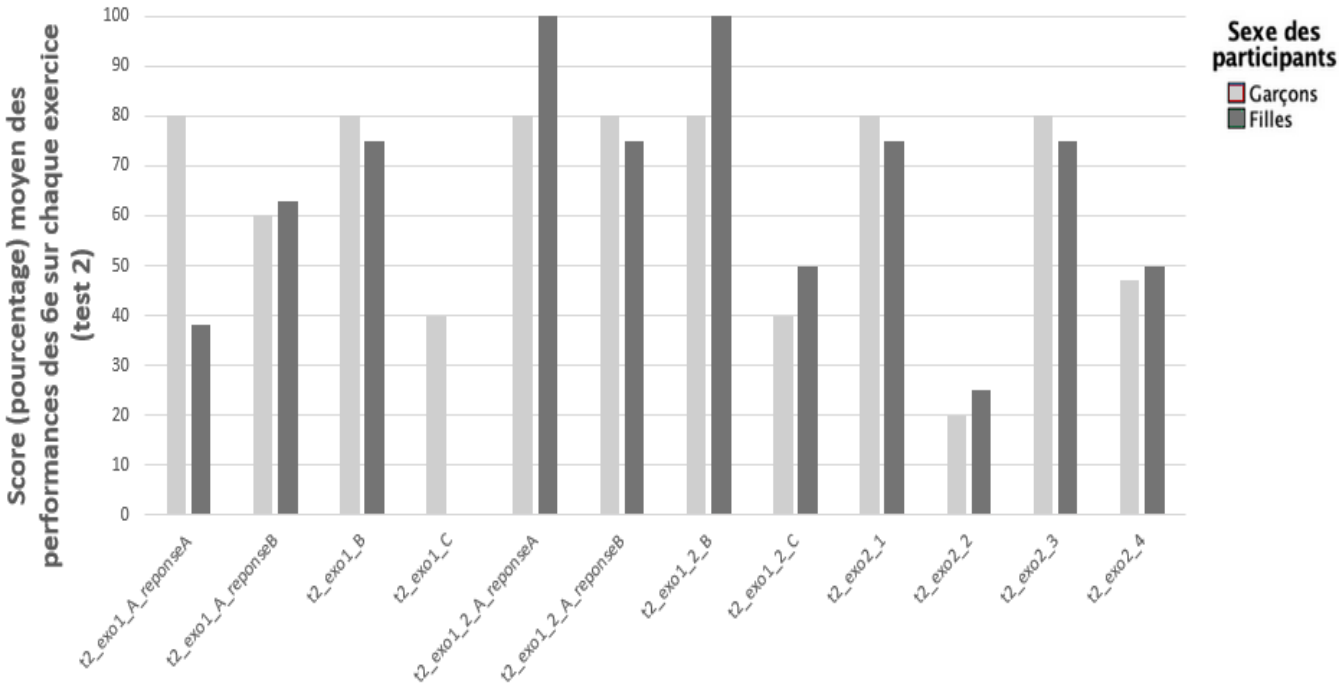
TABLEAU 8.4

Moyenne et erreur standard pour chaque exercice du test 2 des 6e.

Exercice	*M*	*ES*
t2_exo1_A_reponseA	61.10	.16
t2_exo1_A_reponseB	61.10	.16
t2_exo1_B	78.00	.15
t2_exo1_C	22.00	.15
t2_exo1_2_A_reponseA	89.00	.11
t2_exo1_2_A_reponseB	78.00	.15
t2_exo1_2_B	89.00	.11
t2_exo1_2_C	44.00	.18
t2_exo2_1	78.00	.12
t2_exo2_2	22.00	.15
t2_exo2_3	78.00	.15
t2_exo2_4	48.15	.10

FIGURE 8.3.

Étude pilote 2. Score moyen des performances des garçons et des filles de 6e pour chaque exercice du test 2.



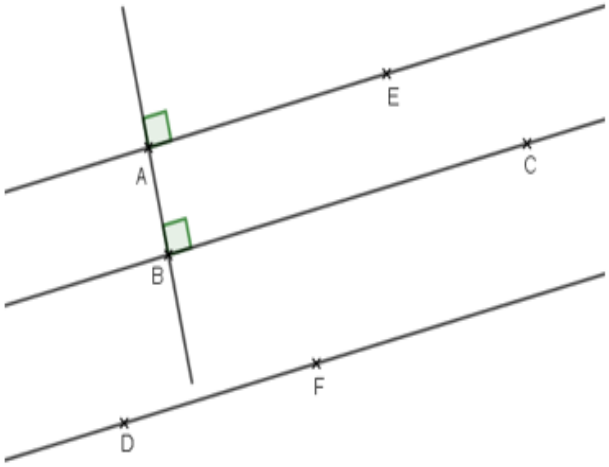
Deux exercices, *t2_exo1_2_A_reponseA* et *t2_exo1_2_B*, ont non seulement été bien réussis

(M=89.00, SD=.11), mais les filles (M=100.00, ES=.00) ont de plus obtenu de meilleurs résultats que les garçons (M=80.00, ES=.45). Nous en avons retiré un des deux, *t2_exo1_2_B*, pour garder un équilibre entre les exercices faciles et plus difficiles. Comme pour le test 1, nous avons rajouté un exercice jugé difficile par les enseignants où il est demandé de justifier la réponse de l'exercice *t2_exo2_1*. Nous avons également rajouté un exercice jugé moyennement difficile. L'exercice supprimé est en Figure 8.4, tandis que les exercices rajoutés sont donnés en Figure 8.5. Le score global des garçons (M=55.45, ES=.19) était similaire à celui des filles (M=52.27, ES=.19). Les filles (M=9.88, ES=1.15) ont effectué le test plus rapidement que les garçons (M=13.83, ES=2.38).

FIGURE 8.4.

Exercice supprimé au test 2 de 6e à cause de sa facilité.

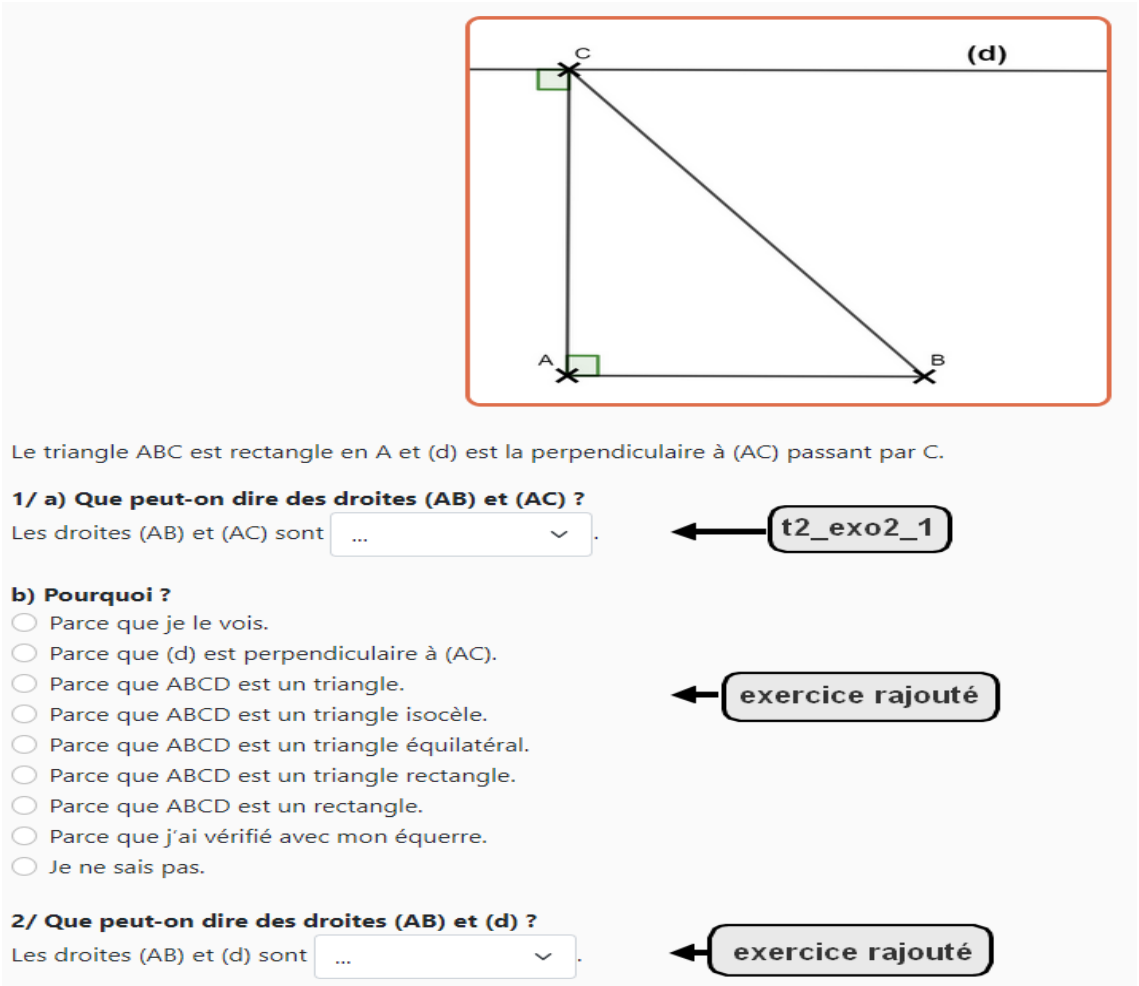
On considère la figure ci-dessous dans laquelle les droites (BC) et (DF) sont parallèles.



b/ Que peut-on en déduire pour les droites (AB) et (DF) ?
Réponse : On peut en déduire que les droites (AB) et (DF) sont ..... (Liste déroulante : sécantes, parallèles, perpendiculaires, confondues)

FIGURE 8.5.

Exercices (jugés difficiles par les enseignants) rajoutés au test 2 de 6e.



8.6.3. Analyses descriptives chez les élèves de 5e

8.6.3.1. Test 1

Le Tableau 8.5 présente la moyenne et l'erreur standard de chaque exercice des 5e pour le test 1, et la Figure 8.6 illustre les différences de résultats à ces exercices entre les filles et les garçons.

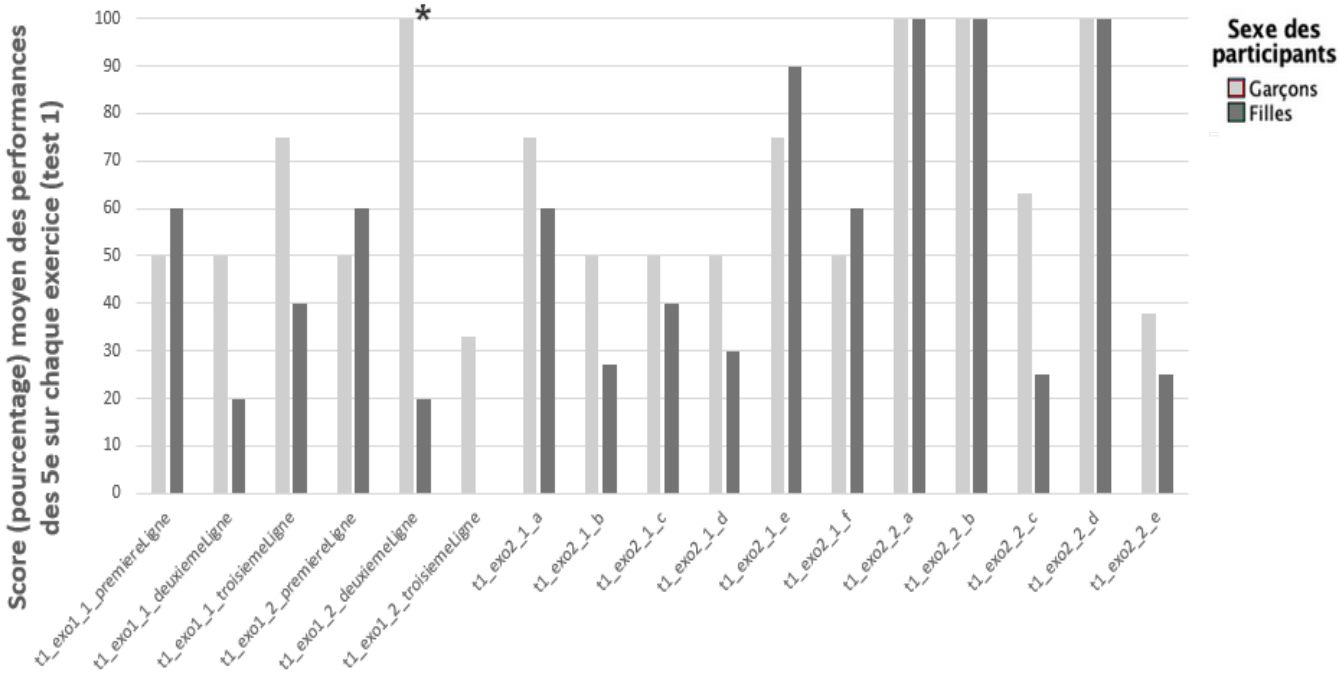
TABLEAU 8.5

Moyenne et erreur standard pour chaque exercice du test 1 des 5e.

Exercice	*M*	*ES*
t1_exo1_1_premiereLigne	56.00	.18
t1_exo1_1_deuxiemeLigne	33.00	.17
t1_exo1_1_troisiemeLigne	56.00	.18
t1_exo1_2_premiereLigne	56.00	.18
t1_exo1_2_deuxiemeLigne	56.00	.18
t1_exo1_2_troisiemeLigne	14.81	.11
t1_exo2_1_a	67.00	.17
t1_exo2_1_b	37.04	.16
t1_exo2_1_c	44.00	.18
t1_exo2_1_d	38.90	.16
t1_exo2_1_e	83.30	.12
t1_exo2_1_f	56.00	.18
t1_exo2_2_a	100.00	.00
t1_exo2_2_b	100.00	.00
t1_exo2_2_c	41.67	.11
t1_exo2_2_d	100.00	.00
t1_exo2_2_e	30.56	.09

FIGURE 8.6.

Étude pilote 2. Score moyen des performances des garçons et des filles de 5e pour chaque exercice du test 1.



Nous avons retiré deux des trois exercices ayant obtenu 100% de réussite, *t1_exo2_2_a* et *t1_exo2_2_b*, pour ne pas avoir un test trop facile. Ces exercices sont donnés en Figure 8.7. Ceci permet également de raccourcir le temps de réponse qui était plus élevé que chez les 6e (chez les filles : M=18.94, ES=4.83 ; chez les garçons : M=17.56, ES=2.01). Les garçons ont obtenu un score global de 64.62% (ES=.24) de bonnes réponses, contre 50.19% (ES=.14) de la part des filles.

FIGURE 8.7.

Exercices supprimés au test 1 de 5e à cause de leur facilité.

2/ Écrire les expressions numériques qui correspondent aux phrases suivantes :
Exemple : La somme de 8 et du quotient de 10 par 2 : 8 + 10 ÷ 2

a/ La somme de 5 et du produit de 3 par 4
Correction : 5+3×4 ou 3×4 + 5 ou 4×3+5

b/ La différence entre le produit de 3 par 8 et 5
Correction : 3 × 8 – 5 ou 8 × 3 - 5

8.6.3.2. Test 2

Concernant le test 2, le Tableau 8.6 montre la moyenne et l'erreur standard pour chaque exercice, et les différences entre les filles et les garçons sont présentées en Figure 8.8.

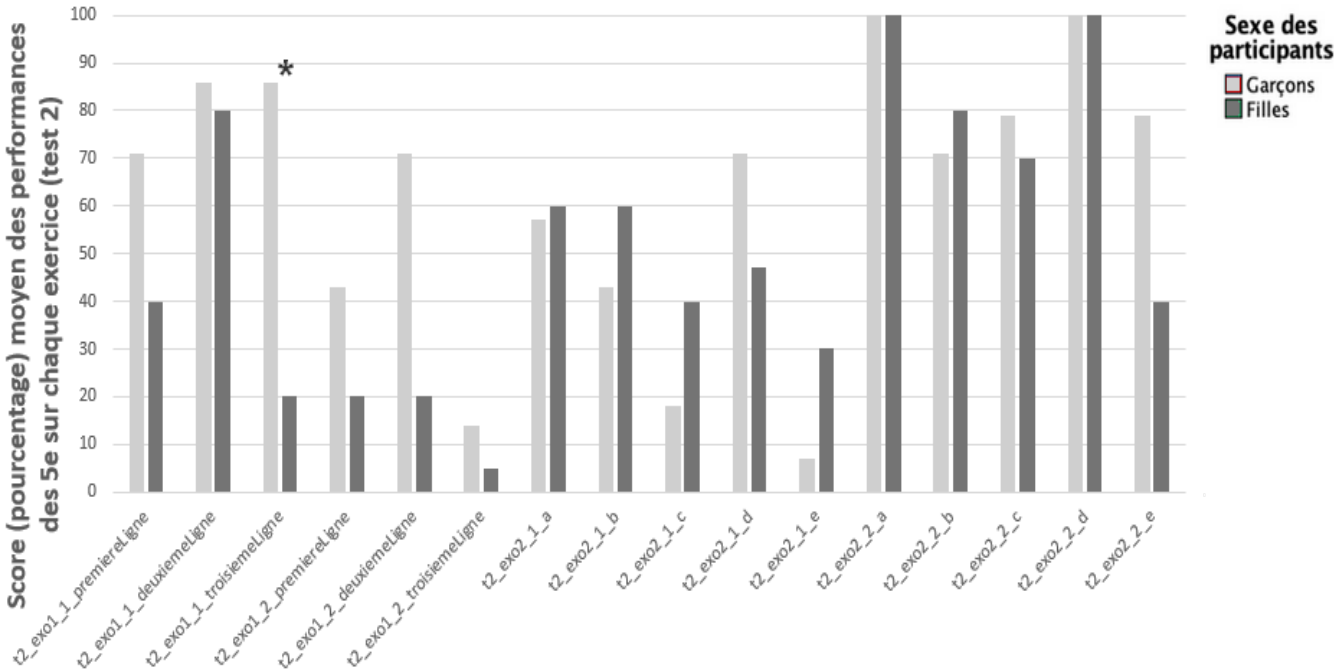
TABLEAU 8.6

Moyenne et erreur standard pour chaque exercice du test 2 des 5e.

Exercice	*M*	*ES*
t2_exo1_1_premiereLigne	58.00	.15
t2_exo1_1_deuxiemeLigne	83.00	.11
t2_exo1_1_troisiemeLigne	58.00	.15
t2_exo1_2_premiereLigne	33.00	.14
t2_exo1_2_deuxiemeLigne	50.00	.15
t2_exo1_2_troisiemeLigne	10.42	.08
t2_exo2_1_a	58.00	.15
t2_exo2_1_b	50.00	.15
t2_exo2_1_c	27.08	.13
t2_exo2_1_e	61.11	.14
t2_exo2_1_f	16.67	.08
t2_exo2_2_a	100.00	.00
t2_exo2_2_b	75.00	.13
t2_exo2_2_c	75.00	.11
t2_exo2_2_d	100.00	.00
t2_exo2_2_e	62.50	.11

FIGURE 8.8.

Score moyen des performances des garçons et des filles de 5e pour chaque exercice du test 2.



Nous avons retiré un des deux exercices ayant obtenu 100% de réussite, *t2_exo2_2_a*. Deux exercices, *t2_exo2_2_b* (M=75.00, ES=.13) et *t2_exo2_2_c* (M=75.00, ES=.11), ont également obtenu un haut score. Nous avons décidé d'en retirer un des deux, celui où les filles (M=80.00, ES=.45) ont mieux réussi que les garçons (M=71.00, ES=.49), *t2_exo2_2_b*. Ceci permet de raccourcir le temps de réponse (chez les filles : M=18.90, ES=2.52 ; chez les garçons : M=18.01, ES=1.78), et également de garder quelques exercices plus faciles afin d'analyser les performances des élèves selon le niveau de difficulté lors de l'étude principale. Les exercices retirés sont donnés en Figure 8.9. Globalement, les filles ont obtenu 51.20% (ES=.16) de bonnes réponses au test 2, contre 60.00% (ES=.12) chez les garçons.

FIGURE 8.9.

Exercices supprimés au test 1 de 5e à cause de leur facilité.

2/ Écrire les expressions numériques qui correspondent aux phrases suivantes :
Exemple : La somme de 8 et du quotient de 10 par 2 : $8 + 10 \div 2$

a/ La différence de 25 et du produit de 3 par 4
Correction : $25 - 3 \times 4$ ou $25 - 4 \times 3$

b/ La somme du quotient de 15 par 3 et de 11
Correction : $(15 \div 3) + 11$ ou $15 \div 3 + 11$ ou $11 + 15 \div 3$ ou $11 + (15 \div 3)$

8.7. Discussion

L'objectif de cette étude pilote était d'examiner la difficulté de chaque exercice composant les quatre tests standardisés de mathématiques, créés en collaboration avec une équipe pédagogique d'enseignants en mathématiques, d'un chargé de mission prioritaire dans les collèges, et d'un inspecteur académique de mathématiques. Ces tests ont été créés pour les utiliser lors de notre étude principale permettant de mesurer l'impact de nos agents virtuels pédagogiques sur la performance en mathématiques des élèves et sur les effets de menace du stéréotype. La menace du stéréotype s'observe généralement sur les exercices les plus compliqués. Cependant, nous voulions préserver un équilibre entre les exercices plus simples et les exercices plus difficiles, afin d'analyser les performances des élèves selon le niveau de difficulté des exercices.

Pour évaluer le niveau de difficulté de tous les exercices des quatre tests standardisés de mathématiques, chaque classe (une classe de 6e et une classe de 5e) a été séparée en deux demi-groupes. Tous les élèves ont visionné une vidéo explicative sur les notions clés de la leçon associée aux tests de mathématiques (le parallélisme pour les 6e, et les priorités opératoires pour les 5e). Les prétests détaillés dans ce chapitre ont montré que certains exercices étaient trop simples, et que les 6e ont terminé rapidement les deux tests de mathématiques. Nous

avons donc supprimé et rajouté des exercices. De plus, nos observations dans les classes ainsi que les retours de l'enseignant ont indiqué que l'outil stylo présent dans les deux tests de 6e a déconcentré les élèves. Cet outil a été initialement conçu pour permettre aux élèves de repasser les droites en couleur, pour qu'ils puissent mieux les identifier. Cependant, malgré l'explication de l'outil dans la vidéo de la leçon, ils ne comprenaient pas à quoi il servait, et pensaient qu'il était obligatoire de repasser toutes les droites. Nous avons donc retiré l'outil stylo pour notre étude principale, et nous avons modifié la leçon de mathématiques en supprimant le conseil incitant à utiliser cet outil ainsi que les explications associées.

Nos analyses ont révélé que globalement, et contrairement aux garçons, les filles ont produit des contre-performances par rapport à leur moyenne trimestrielle en mathématiques. Bien que l'on ne retrouve pas de différence filles/garçons dans les performances aux tests standardisés de mathématiques, il est fort probable que ce résultat soit en partie dû à la menace du stéréotype. En effet, en France, les écarts entre filles et garçons en mathématiques se réduisent de plus en plus, mais cela est dû à une baisse de performance chez les garçons et non pas à une éventuelle baisse des effets de menace du stéréotype (FUMEL et al., 2016).

Compte tenu de ces résultats, nous avons rajouté un exercice jugé difficile par l'équipe pédagogique pour le test 1 des 6e, ce qui donne 8 exercices. Pour le test 2, nous avons retiré un exercice facile et ajouté deux exercices, jugés comme difficile et moyennement difficile par l'équipe pédagogique, donnant lieu à 13 exercices. Concernant les 5e, nous avons retiré 2 exercices faciles aux deux tests. Le test 1 contient 15 exercices, et le test 2 en contient 14. Tous les exercices prétestés dans ce chapitre ainsi que les exercices finaux sont présentés en Annexe F (test 1 des 6e), G (test 2 des 6e), H (test 1 des 5e), et I (test 2 des 5e).

8.8. Points clés du chapitre

- Création et élaboration de 4 tests standardisés de mathématiques (2 pour les 6e, 2 pour les 5e) et de 2 leçons de mathématiques pour chaque niveau (6e : le parallélisme, 5e : les priorités opératoires).
- Collaboration avec une équipe pédagogique en mathématiques : cinq enseignants de mathématiques d'un collège en REP+, un chargé d'éducation en zone prioritaire (également enseignant de mathématiques au collège), et un inspecteur académique en mathématiques.
- Évaluation de la difficulté de chaque exercice auprès d'une classe de 6e et d'une classe de 5e.
- Pas de différence de performance entre filles et garçons.
- Les filles ont significativement moins bien réussi les exercices comparativement à leur moyenne trimestrielle en mathématiques, contrairement aux garçons ayant obtenu les mêmes résultats.
- Ajustement de la difficulté des tests en ajoutant des exercices (jugés difficiles par l'équipe pédagogique) et en supprimant des exercices trop faciles.

9. Étude de l'impact des agents virtuels pédagogiques sur la performance des élèves en mathématiques

Sommaire

9.1. Objectifs . . . 217
9.2. Méthode . . . 222
9.2.1. Participants . . . 222
9.2.2. Procédure . . . 222
9.2.3. Mesures . . . 224
9.2.3.1. Questionnaires d'auto-évaluation en mathématiques . . . 224
9.2.3.2. Test des Associations Implicites genre-mathématiques-lecture 225
9.2.3.3. Performances aux tests standardisés de mathématiques . . . 230
9.2.3.4. Moyenne trimestrielle en mathématiques . . . 230
9.3. Résultats . . . 231
9.3.1. Analyses préliminaires . . . 231
9.3.1.1. Statistiques descriptives . . . 231
9.3.1.2. Analyses sur les effets potentiels du type de notation en classe 233
9.3.1.3. Analyses sur les biais de répartition des élèves dans les conditions expérimentales en fonction de leur niveau initial en mathématiques . . . 240
9.3.2. Analyses centrales . . . 241
9.3.2.1. Analyses au temps 1 . . . 241
9.3.2.2. Analyses au temps 2 . . . 243
9.3.2.3. Analyses de l'effet modérateur du TAI . . . 246
9.4. Discussion . . . 247
9.5. Points clés du chapitre . . . 253

9.1. Objectifs

Ce chapitre détaille l'étude principale de cette thèse : **l'évaluation de l'impact de nos agents virtuels pédagogiques sur les performances des élèves, et plus particulièrement des filles, en mathématiques**. Comme nous l'avons vu (Chapitre 2), plusieurs recherches en cognition sociale ont montré que l'effet de menace du stéréotype contribue à diminuer les performances des filles en mathématiques, et qu'il est possible d'en réduire les effets négatifs grâce, notamment, à l'exposition à des modèles de réussite en mathématique dès lors que la réussite de ces modèles est perçue comme atteignable. L'objectif de l'étude présentée dans ce chapitre est d'améliorer les performances en mathématiques des élèves du secondaire grâce à des agents virtuels pédagogiques créés pour réduire les effets de menace du stéréotype dans ce domaine. Plus précisément, il s'agit d'évaluer l'efficacité d'agents virtuels pédagogiques conçus pour, d'une part, incarner des modèles de réussite en mathématiques et d'autre part, adopter des comportements propices à l'établissement d'un rapport virtuel avec les élèves afin de favoriser leur identification à la réussite des agents. C'est la première fois qu'une étude est menée pour tester l'efficacité d'agents virtuels pédagogiques spécifiquement développés pour réduire les effets de menace du stéréotype en mathématiques auprès d'enfants en situation de classe quasi-ordinaire. Les résultats que nous attendions dans le cadre de cette thèse sont une meilleure compréhension des profils des modèles de réussite pour réduire les effets de menace du stéréotype. Nous visions également à mieux comprendre la façon dont la recherche en informatique peut produire des outils pouvant être utilisés pour lutter contre ce phénomène lors des apprentissages. Les résultats permettraient de mesurer l'amélioration des performances des élèves à l'aide d'agents virtuels pédagogiques incarnant des modèles de réussite.

Comme cela est décrit dans nos chapitres précédents (Chapitres 6 et 7), nos agents virtuels ont été créés à partir des études en cognition sociale qui ont identifié les caractéristiques nécessaires pour qu'un individu représente un modèle de réussite efficace en fonction des cibles visées. Ainsi, nous avons créé deux agents virtuels pédagogiques de genre différent, une fille et un garçon, dont l'apparence physique est légèrement plus âgée que les élèves, et qui se présentent comme étant au lycée. Une évaluation de l'âge a permis de vérifier que les élèves percevaient bien les agents comme étant légèrement plus âgés qu'eux (Section 6.2.3). Ensuite, outre l'apparence physique, un modèle de réussite doit montrer qu'il est compétent dans le domaine en question (ici les mathématiques), et sa réussite doit être perçue comme étant à la portée de l'élève. Les études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016) ont montré que cette identification à la réussite d'un modèle humain est possible dès lors que la réussite est expliquée par l'effort et le travail. Quant aux études en informatique sur les agents virtuels pédagogiques, elles ont montré qu'un agent virtuel de ce type est d'autant plus efficace s'il induit, par ses comportements verbaux et non verbaux, un rapport virtuel avec l'apprenant. Conformément à ces deux ensembles de littérature (Chapitres 2 et 3), chacun de nos deux agents virtuels, féminin et masculin, ont adopté l'un des trois types de modèles comportementaux suivants (voir Chapitre 6 pour plus de détails sur leurs modèles computationnels). Le premier modèle de comportements, dont nous supposons qu'il devrait être le plus efficace pour réduire les effets de menace du stéréotype, consiste pour l'agent à se présenter comme en réussite en mathématique, à expliquer sa réussite par ses efforts et son travail, et à adopter des comportements verbaux et non verbaux tout au long de l'interaction afin d'instaurer un rapport virtuel avec l'élève (Annexe D pour tous les dialogues). Nous ap-

pelons cet agent "MER" pour "**Modèle avec Explications et Rapport**". Le second modèle de comportement pour l'agent consiste à se présenter comme ayant réussi en mathématiques et à expliquer sa réussite par ses efforts et son travail, mais sans manifester de comportements de rapport virtuel. Cet agent est appelé "$ME\overline{R}$" pour "**Modèle avec Explications et Sans Rapport**". Enfin, dans le troisième modèle de comportements, l'agent ne donne aucune explication de sa réussite et ne manifeste aucun comportement de rapport virtuel (Annexe E pour les dialogues des deux derniers types d'agents). Ce dernier agent est appelé "$M\overline{ER}$" pour "**Modèle Sans Explications et Sans Rapport**".

Les évaluations de nos agents virtuels pédagogiques animés (Chapitre 7) ont montré que les élèves ont trouvé les agents très compétents en mathématiques, se sont identifiés à eux et à leur réussite dans cette matière. Ces perceptions ont été renforcées par les comportements de rapport virtuel de l'agent. Le rapport virtuel a donc permis aux agents d'incarner de meilleurs modèles de réussite en mathématiques. De plus, les élèves (filles comme garçons) se sont plus identifiés à l'agent virtuel féminin qu'à l'agent virtuel masculin. Comme l'ont montré plusieurs études sur le sujet, on s'attendait à ce que ce soit uniquement les modèles de réussite incarnés par des filles/femmes qui pouvaient réduire les effets de menace du stéréotype. Or, d'autres études ont montré que finalement, le genre du modèle n'avait pas d'importance dès lors qu'il explique sa réussite en mathématiques par des efforts et du travail (BAGÈS & MARTINOT, 2011 ; BAGÈS et al., 2016). À noter cependant que, contrairement aux premières études mentionnées, il n'était pas possible de voir le modèle de réussite dans celles de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016), puisqu'il s'agissait de lire les histoires personnelles des modèles. Nous n'avons donc pas d'hypothèse tranchée sur la question du genre de l'agent virtuel pédagogique. Les résultats du chapitre 7 montrant que les élèves se sont plus identifiés à l'agent virtuel féminin qu'à l'agent virtuel masculin tendent cependant à penser que l'agent virtuel féminin serait plus efficace pour réduire les effets de menace du stéréotype. L'hypothèse principale de cette étude est qu'**un agent virtuel pédagogique peut réduire les effets de menace du stéréotype, et donc améliorer les performances des filles -sans pour autant baisser celles des garçons- en incarnant un modèle de réussite capable d'expliquer sa réussite en mathématiques et capable de créer un rapport virtuel avec les élèves**.

La présente étude a été réalisée auprès d'élèves de 6e et de 5e testés dans leur établissement scolaire, en situation de classe quasi-ordinaire. En effet, bien que l'expérience se soit déroulée en classe et non en laboratoire, la présence de l'expérimentatrice ainsi que l'expérimentation en tant que telle (e.g., tests standardisés et agents virtuels) ne correspondaient pas au standard auquel les élèves sont habitués. L'effet des différents agents virtuels a été évalué sur la performance des élèves à un test standardisé de mathématiques, et un questionnaire a permis de mesurer leur sentiment de réussite en mathématiques ainsi que l'importance personnelle accordée à ce domaine de compétence. Également, une mesure implicite des stéréotypes de genre en mathématiques et en lecture (Test des Associations Implicites, TAI - CVENCEK et al., 2011) a été utilisée pour estimer la force d'ancrage de ces stéréotypes en mémoire chez les élèves de cette tranche d'âge, et examiner si ces automatismes stéréotypés sont susceptibles de modérer l'effet des agents virtuels. Le TAI de CVENCEK et al. (2011) n'est pas tout à fait le même que celui pour les adultes de GREENWALD et al. (2003). À ce jour et à notre connaissance, aucune version française du TAI pour enfants/pré-adolescents de CVENCEK et al. (2011) n'a été créée. Nous avons donc codé le TAI genre-mathématiques-lecture en français pour enfants et pré-adolescents, en HTML, JavaScript, et PHP pour la version web (disponible sur

https://github.com/ArmandoMarjorie/French_child_IAT), et en C# pour l'application Unity sur tablette.

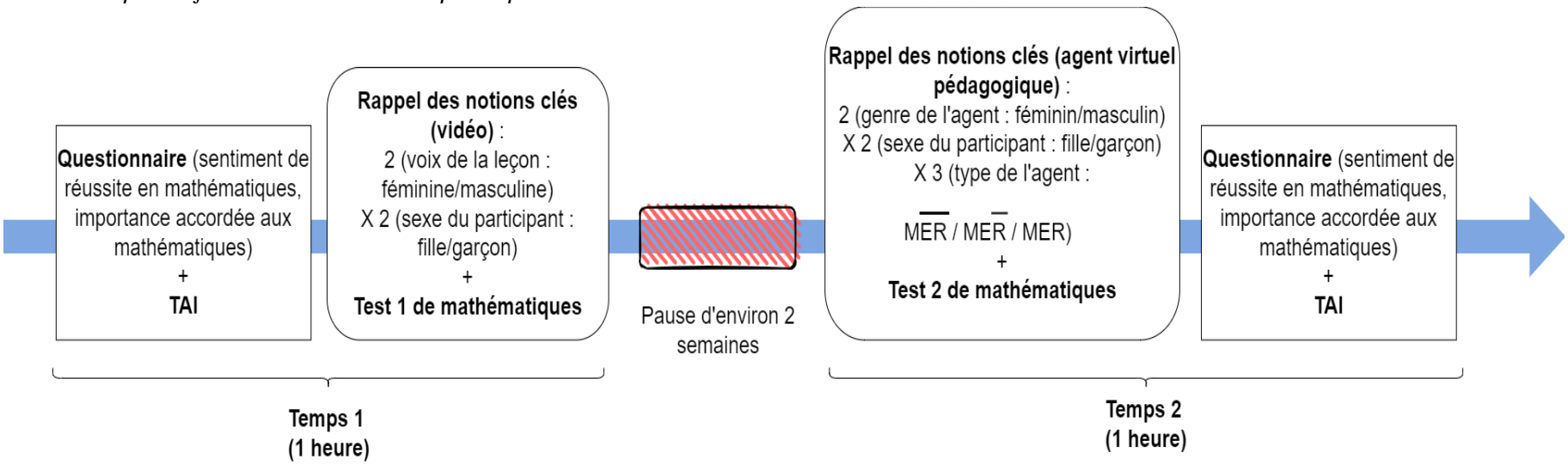
Les questionnaires, le TAI, et les tests de mathématiques ont été administrés en deux phases (Figure 9.1 pour une visualisation schématique de la procédure). La première phase se déroulait sans la présence d'agent virtuel : seule une voix audio pré-enregistrée et générée par Text-to-Speech (la même que celle utilisée par les agents) expliquait la leçon de mathématiques et introduisait le test. Cette phase avait pour objectif de familiariser les élèves avec les conditions d'expérimentations et l'utilisation du numérique et d'obtenir une *baseline* pour les auto-évaluations et la stéréotypie implicite. Nous avons vérifié que les leçons de mathématiques abordées ont bien été enseignées précédemment dans le programme scolaire dans toutes les classes participantes. Cependant, des disparités existaient forcément entre ces classes (pédagogie de l'enseignant, écart de temps entre la passation de notre étude et l'enseignement de la leçon dans le programme scolaire). La leçon a été expliquée pour que tous les élèves puissent se rappeler des notions de la même façon, et mettre le plus possible tous les participants au même niveau.

La deuxième phase se déroulait en présence des agents virtuels pédagogiques, qui se présentaient et expliquaient la leçon de mathématique (à la place de la vidéo présentée lors du premier temps). Le test de mathématique proposé portait sur les mêmes notions que le test passé dans la phase précédente, mais les exercices étaient différents pour limiter l'effet d'apprentissage et ne pas biaiser l'évaluation au second temps. Nous avons donc analysé les résultats obtenus lors de ces deux temps séparément. Comme expliqué ci-dessus, nous avons formulé l'hypothèse que l'agent virtuel pédagogique expliquant sa réussite en mathématiques par le travail et avec des comportements de rapport virtuel (l'agent MER) permettrait aux élèves, et plus particulièrement aux filles, d'obtenir les meilleurs résultats. Nous n'avons pas d'hypothèse particulièrement tranchée sur le genre de l'agent, car bien que plusieurs études aient montré que ce sont les modèles de réussite incarnés par des filles ou des femmes qui permettaient de réduire les effets de menace du stéréotype, d'autres études ont montré que finalement, le genre du modèle de réussite n'avait pas d'importance à partir du moment où celui-ci expliquait sa réussite par du travail et des efforts. Nous avons également formulé l'hypothèse que le rapport virtuel permettrait d'améliorer le sentiment de réussite et l'importance aux mathématiques des élèves. Plus particulièrement, nous pensions que l'agent féminin permettrait aux filles de se sentir plus en réussite en mathématiques et d'y accorder plus d'importance. Enfin, nous nous sommes demandé si les scores de stéréotypie implicite chez les filles et les garçons modéraient l'influence de l'agent virtuel sur leurs auto-évaluations et/ou leurs performances en mathématiques. En effet, deux résultats contradictoires existent dans la littérature concernant l'effet modérateur de la stéréotypie implicite de la menace du stéréotype sur les performances en mathématiques des femmes, pouvant être expliqués par les différences de conditions expérimentales. KIEFER et SEKAQUAPTEWA (2007) ont ainsi montré que préciser que les compétences en mathématiques n'étaient pas évaluées lors du test suffisait à augmenter les résultats des femmes avec de faibles stéréotypies implicites, mais pas pour celles ayant obtenu un haut score au TAI sciences-genre. À l'inverse, ce sont les femmes avec des stéréotypies implicites élevées qui ont bénéficié de la condition de menace réduite dans l'étude de FRANCESCHINI et al. (2014), où il était précisé que les femmes avaient obtenu de meilleurs résultats que les hommes dans le test. La question reste entière sur une population comme la nôtre (des élèves de 6e et de 5e). Notre induction expérimentale est

tout de même plus proche de celle de FRANCESCHINI et al. (2014). Nous avons donc formulé l'hypothèse que l'agent féminin expliquant sa réussite en mathématiques par le travail, et exprimant des comportements de rapport virtuel, permettrait d'améliorer les performances des filles ayant un score de stéréotypie implicite élevé, ainsi que leurs auto-évaluations.

FIGURE 9.1.

Schéma explicatif des tâches de l'étude principale.



9.2. Méthode

9.2.1. Participants

Les enseignants et enseignantes de mathématiques de 23 classes ont accepté de participer à l'étude, dont 11 classes de 6e et 12 classes de 5e, provenant de 4 collèges différents. Un total de 574 élèves étaient inscrits sur les listes des classes, dont 143 garçons de 6e (M_{age}=11.32, ET=.05), 166 garçons de 5e (M_{age}=12.39, ET=.04), 136 filles de 6e (M_{age}=11.35, ET=.05), et 129 filles de 5e (M_{age}=12.22, ET=.05). Les classes ont été réparties dans les différentes conditions expérimentales afin de tendre vers un nombre relativement équivalent de filles et de garçons dans chaque condition, avec des moyennes trimestrielles comparables en mathématiques. Parmi ces 574 élèves, 63 élèves présentaient des difficultés d'apprentissage parmi lesquelles nous trouvons des élèves allophones suivant des cours de FLE (Français Langue Étrangère), des élèves à profil ULIS (Unités Localisées pour l'Inclusion Scolaire), des élèves à profil PAP (Plan d'Accompagnement Personnalisé), des élèves bénéficiant de l'aide d'un AESH (Accompagnant des Élèves en Situation de Handicap) ou d'un AVS (Auxiliaire de Vie Scolaire), des élèves à profil PPS (Projet Personnalisé de Scolarisation), des élèves à profil UPE2A (Unités Pédagogiques pour Élèves Allophones Arrivants), des élèves ne sachant pas lire, et des élèves absents des cours. Ces élèves ont été retirés des analyses, mais ils ont tout de même participé à l'étude, dans la même salle et en même temps que leurs autres camarades de classe, et ce pour deux raisons. Il s'agissait d'une part, d'éviter de stigmatiser ces élèves et d'autre part, de maintenir la pression évaluative associée au test de mathématiques (les notes vont compter dans la moyenne) qui est une condition essentielle pour observer les effets de menace du stéréotype.

Un total de 511 élèves étaient éligibles pour participer à l'étude. Parmi ces 511 élèves, 443 étaient présents pour la passation au temps 1 (soit 86.7%). Parmi eux, 32 élèves ont été exclus des traitements, soit pour mauvais comportements pendant les passations, soit pour refus parental de participation, soit en raison d'un bug informatique (perte de signal wifi). Finalement, le Tableau 9.1 montre un total de 411 élèves (212 filles et 199 garçons) se répartissant dans les 4 conditions expérimentales, pour le temps 1. Sur les 411 élèves retenus pour les analyses au temps 1, 375 ont été présents au temps 2 (soit 91.2%). Le Tableau 9.2 montre la répartition de ces élèves dans les 12 conditions expérimentales du temps 2.

TABLEAU 9.1

Nombre de participants par conditions expérimentales au temps 1, après avoir appliqué les critères d'exclusions.

Sexe du participant	Voix de la leçon	
	Masculine	Féminine
Garçon	104	108
Fille	100	99

9.2.2. Procédure

Les élèves de chaque classe se sont installés dans la salle informatique de leur collège, ou bien dans leur salle de classe habituelle, avec leur enseignant de mathématiques ou un

TABLEAU 9.2

Nombre de participants par conditions expérimentales présents à la fois aux temps 1 et 2, après avoir appliqué les critères d'exclusions.

Genre de l'agent	Sexe du participant	Type de l'agent		
		$M\overline{ER}$	$ME\overline{R}$	MER
Masculin	Garçon	34	30	32
	Fille	31	30	25
Féminin	Garçon	34	35	32
	Fille	30	29	33

personnel de l'établissement (e.g., CPE, assistants d'éducation). Si l'établissement ne disposait pas d'ordinateurs compatibles avec Unity WebGL, une tablette (Ipad modèle MR7F2NF/A, iOS 11.4) était prêtée à chaque élève. Les élèves étaient tous équipés de casque audio fourni par Ampiric (MIRAULT et al., 2023). Les classes d'un même établissement et du même niveau (6e/5e) ont passé l'expérimentation en parallèle afin d'éviter que les élèves ne parlent entre eux de l'étude avant que tous ne l'aient passée. Les élèves dont les parents ont refusé la participation de leur enfant ont passé les tests de mathématiques à l'écrit, dans la même salle que leurs camarades, afin de ne pas laisser croire aux autres élèves que les évaluations ne compteraient pas. Ces élèves n'ont pas été pris en compte dans les analyses. Cette étude s'est déroulée de janvier à mai.

L'expérimentatrice est passée dans chaque classe en début de séance pour expliquer aux élèves les raisons de sa présence dans la classe et la procédure à suivre. Elle s'est présentée en tant que personne travaillant pour les collèges : « j'ai été contactée par les chefs d'établissement afin de faire passer les mêmes bilans de compétences en mathématiques dans les collèges de la région ». Il a été spécifié aux élèves que les tests allaient se faire sur ordinateurs (ou tablettes) et que les résultats compteraient dans leur moyenne trimestrielle de mathématiques. L'expérimentatrice a expliqué qu'avec son équipe, elle avait conçu les applications qui allaient être utilisées sur les tablettes/ordinateurs, mais que les tests de mathématiques avaient été conçus par les enseignants et les inspecteurs académiques. L'expérimentatrice a précisé qu'il y aurait deux évaluations à quelques semaines d'intervalle, sans préciser les notions abordées afin que les élèves ne s'entraînent pas entre temps, afin d'obtenir un bilan approfondi sur les niveaux en mathématiques des élèves. Lors de chacune des deux séances, les élèves pouvaient poser des questions en levant le doigt s'ils ne comprenaient pas un exercice ou s'ils rencontraient des difficultés à naviguer sur les applications. Ils n'avaient pas le droit à la calculatrice, mais pouvait prendre une feuille de brouillon. Les élèves ayant terminé avant la fin de l'heure étaient invités à rester silencieux et à faire leurs devoirs, ou bien à continuer leur avancée de leurs cours de mathématiques sur le site de Studio Code (application en ligne pour apprendre les bases de la programmation, utilisée dans les classes). À la fin de chaque séance, les élèves étaient remerciés.

Temps 1. Il était expliqué aux élèves qu'ils allaient passer lors de cette première séance un test de mathématiques de 30 minutes, et qu'il y avait d'autres tâches à effectuer avant. Les élèves étaient tout d'abord invités à répondre à un questionnaire dont l'objectif était de re-

censer les avis des élèves de collèges sur les mathématiques. L'expérimentatrice insistait sur le caractère totalement anonyme de leurs réponses à ce questionnaire, que leur enseignant n'aurait pas accès à leurs opinions. Elle précisait qu'il n'y avait ni bonne ni mauvaise réponse au questionnaire, et que l'important était de donner leur avis le plus honnêtement possible. Ensuite, le Test des Associations Implicites (TAI) que nous avons recodé était présenté aux élèves comme un test de réflexe et de rapidité pour savoir si les élèves savaient bien utiliser un ordinateur/une tablette. L'expérimentatrice a donné plusieurs exemples de configuration du TAI au tableau. Enfin, les élèves étaient informés que l'évaluation en mathématiques comprenait deux temps : un premier temps consacré au visionnage d'une vidéo d'environ 4 minutes où une voix off (féminine ou masculine) expliquait une leçon sur des notions de mathématiques, et un deuxième temps qui correspondait au test de mathématiques sur les notions abordées dans la leçon. Il était recommandé aux élèves d'être très attentifs à la leçon afin de bien réussir le test de mathématiques. Les élèves pouvaient poser des questions sur le déroulement de la séance, ou s'ils n'avaient pas compris une tâche. Une fois les explications terminées, les élèves mettaient leur casque et commençaient la séance, avec dans l'ordre, le questionnaire, le TAI, la leçon et le test de mathématiques.

Temps 2. Les élèves étaient informés que cette seconde et dernière séance était comparable à la précédente, avec quelques différences, et que l'objectif du deuxième test était d'évaluer leurs progrès en mathématiques. Comme au temps 1, il était expliqué aux élèves qu'ils allaient passer un test de mathématiques de 30 minutes, qui serait noté et compterait dans la moyenne. Cette fois-ci, ils devaient commencer par la leçon et le test de mathématiques, puis répondre au questionnaire portant sur leurs opinions vis-à-vis des mathématiques, et enfin passer le TAI. L'expérimentatrice fournissait les mêmes explications que lors du temps 1 pour le questionnaire et le TAI, d'une part pour que les élèves absents au temps 1 puissent comprendre, et d'autre part parce qu'il est probable que les élèves ayant déjà entendu les explications les aient oubliées. L'expérimentatrice expliquait ensuite qu'avant de passer le test de mathématiques, ils allaient recevoir une leçon de mathématiques dispensée par un personnage virtuel en 3D créé à partir d'un vrai élève au lycée, de son apparence physique jusqu'à son histoire personnelle. Comme pour le temps 1, il était bien précisé que la leçon de mathématiques portait sur des notions qui seraient évaluées par le test, et qu'il fallait donc être très attentif pour réussir le test de mathématiques. L'expérimentatrice répondait aux questions éventuelles, puis invitait les élèves à mettre leur casque et la séance pouvait alors débuter.

9.2.3. Mesures

Cette section présente les différentes variables de l'étude, lesquelles ont toutes été utilisées lors des deux temps de mesure. Nous présentons donc ci-dessous le contenu des questionnaires, le test des associations implicites, les performances aux tests de mathématiques, et les évaluations trimestrielles des élèves en mathématiques.

9.2.3.1. Questionnaires d'auto-évaluation en mathématiques

Le questionnaire, très court, permettait de recueillir le sentiment de réussite des élèves en mathématiques et leur importance accordée à ce domaine de compétences. Les élèves étaient invités à donner leurs réponses sur une échelle de Likert en 5 points (de 1 : "Pas du

tout d'accord" à 5 : "Tout à fait d'accord"), en sélectionnant le chiffre correspondant le mieux à leur avis. Un score moyen a été calculé pour chacune de ces mesures.

Sentiment de réussite en mathématiques. Les élèves indiquaient leur sentiment de réussite en mathématiques à l'aide des 2 items suivants : (1) "Je réussis bien en mathématiques", (2) "Je suis capable d'avoir des bonnes notes en maths". Les deux items corrélant positivement de manière significative (temps 1 : $r(411)=.52$, $p<.001$; temps 2 : $r(326)=.49$, $p<.001$), un score moyen a été calculé (un score élevé indiquant un sentiment élevé de réussite en mathématiques en général).

Importance accordée aux mathématiques. Les élèves répondaient aux 3 items suivants pour évaluer dans quelle mesure les mathématiques étaient importantes pour eux personnellement : (1) "Je considère qu'il est important pour moi de réussir en maths", (2) "C'est important, pour moi, d'avoir des bonnes notes en maths", (3) "J'ai envie de faire de mon mieux en maths". L'alpha de Cronbach indiquait une bonne cohérence interne ($\alpha=.70$ aux deux temps), permettant de moyenner les trois items (un score élevé indiquait une grande importance accordée aux mathématiques).

9.2.3.2. Test des Associations Implicites genre-mathématiques-lecture

Le Test des Associations Implicites (TAI) "sciences et genre" est une tâche informatisée de catégorisation de mots, permettant de déterminer si les individus ont tendance à associer plus fréquemment les hommes aux sciences et les femmes aux lettres plutôt que l'inverse (GREENWALD et al., 2003). Le TAI est un outil permettant de faire le lien entre les processus psychologiques et l'influence de notre environnement sur ceux-ci. Des associations implicites entre des concepts se développent à travers une multitude d'expériences vécues dans la vie quotidienne. Ce test ne révèle donc pas nécessairement les opinions véritables d'un individu. Le résultat du TAI peut même contredire les croyances explicites d'une personne. Lors de la tâche de catégorisation du TAI, les temps de réponse, mesurés en millisecondes, sont enregistrés pour estimer la force des associations implicites entre différents concepts : certaines informations sont fortement associées entre elles, de par leur utilisation répétée au quotidien, d'autres moins, d'autres pas du tout. Il est intéressant de noter que la version enfant du TAI utilisée par CVENCEK et al. (2011), avec des mots prononcés lorsqu'ils apparaissent à l'écran, a montré que les enfants âgés de 6 à 10 ans ont des associations plus fortes entre mathématiques-garçons qu'entre mathématiques-filles, ce qui indique que les associations stéréotypées sont également présentes chez les enfants.

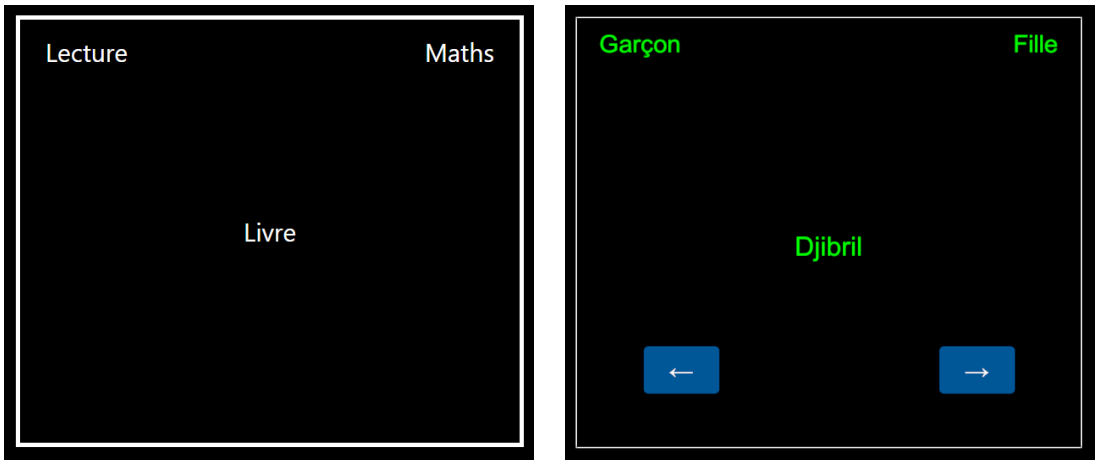
Cependant, peu d'études ont utilisé le TAI avec des enfants. À notre connaissance, aucune version française du TAI de CVENCEK et al. (2011) n'est disponible. Nous avons donc adapté le TAI pour les enfants, en français. D'une part, il ne s'agit pas du TAI sciences et genre habituellement passé chez les adultes, mais du TAI genre-mathématiques-lecture. Le contexte de l'étude de CVENCEK et al. (2011) étant similaire au notre, l'explication du choix de la lecture en tant que contraste aux mathématiques s'explique de la même manière que les auteurs : l'enseignement de la lecture et des mathématiques est obligatoire au collège, et les tests standardisés tels que le brevet des collèges comportent un test en lecture et en mathématiques. D'autre part, les mots à catégoriser sont énoncés à l'aide d'une voix pré-enregistrée pour éviter les biais

de lecture. Nous l'avons codé en HTML, JavaScript et PHP pour la version web (disponible sur https://github.com/ArmandoMarjorie/French_child_IAT), et en C# pour l'application Unity sur tablette. Nous avons utilisé ce TAI genre-mathématiques-lecture pour enfants afin de mesurer la force des associations stéréotypiques (garçon avec mathématiques, fille avec lecture) ou contre-stéréotypiques (fille avec mathématiques, garçon avec lecture). Pour cela, ce TAI utilise 4 catégories : garçon, fille, maths, et lecture. Différents mots et prénoms sont à catégoriser dans chacune des catégories, comme "Lucas" pour "garçon", "Nadia" pour "fille", "addition" pour "maths", et "livre" pour "lecture". Toutes les catégories ainsi que leurs mots associés sont listées en Annexe K. Le TAI est une tâche intrasujets impliquant deux phases (une phase d'entraînement et une phase de test), dans lesquelles les appariements de ces 4 catégories varient. Les mots représentant chacune des catégories sont présentés, un à un, au centre de l'écran de l'ordinateur (ou de la tablette), et les participants doivent les classer dans leur catégorie d'appartenance, en appuyant sur l'une des deux touches du clavier E ou I, ou bien directement sur les flèches sur tablette (Figure 9.2).

FIGURE 9.2.

Exemples du TAI genre-mathématiques-lecture où l'on doit classer les mots avec les touches E et I du clavier (a) ou avec les flèches sur l'écran tactile de la tablette (b).

(a) (b)

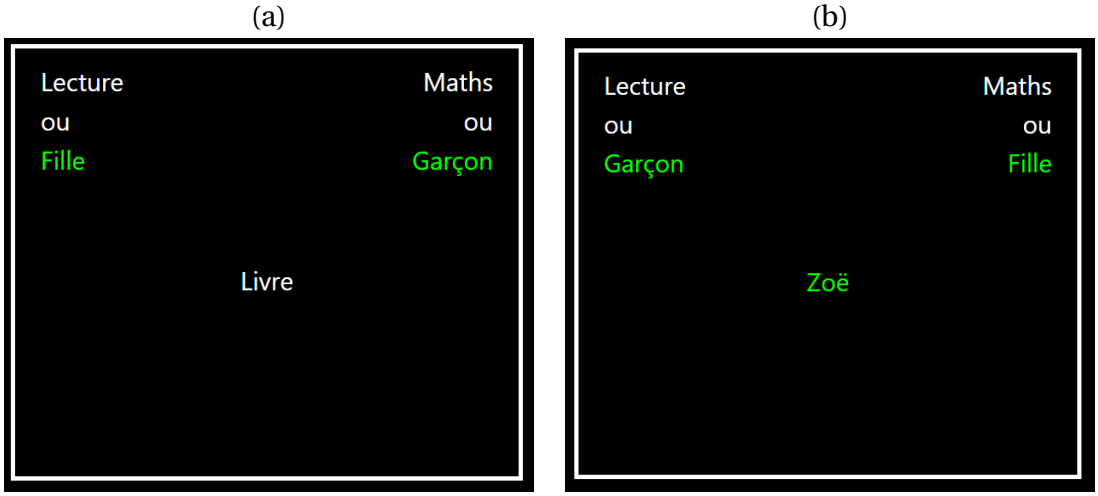


Dans la condition congruente, les catégories "maths" et "garçon" apparaissent ensemble du même côté de l'écran (par exemple en haut à droite ; cf. Figure 9.3 (a)) et partagent donc la même touche de réponse sur le clavier pour catégoriser le mot au centre de l'écran (la touche "I"). Dans cette même condition, les catégories "fille" et "lecture" apparaissent de l'autre côté de l'écran (en haut à gauche) et sont associées à la touche de réponse "E" sur le clavier, pour catégoriser le mot correspondant à cet ensemble de catégories. Dans la condition incongruente, ce sont les catégories "maths" et "fille" qui apparaissent d'une part (par exemple en haut à droite ; cf. Figure 9.3 (b)), et "lecture" et "garçon" qui apparaissent de l'autre côté de l'écran. L'ordre de ces conditions est randomisé. La différence des temps moyens de réponse de catégorisation entre les deux conditions est un indicateur de la force d'association entre un genre (garçon ou fille) et une matière (mathématiques ou lecture). Le score du TAI, compris entre -2 et 2, peut être positif (tendance stéréotypique) ou négatif (tendance contre-stéréotypique).

Plus la valeur absolue du score est élevée, plus la force des associations est importante.

FIGURE 9.3.

Exemples du TAI genre-mathématiques-lecture de la condition congruente (a) et incongruente (b) avec le stéréotype.



Le calcul du score du TAI est donné en pseudo-code dans l'Algorithme 1. Nous prenons pour exemple la configuration où les *blocks* congruents avec le stéréotype ont été passés en premier. $B3$ et $B4$ correspondent ainsi aux *blocks* où les catégories sont congruentes avec le stéréotype, alors que les catégories des *blocks* $B6$ et $B7$ sont incongruentes avec le stéréotype. Seule la phase de test, comprenant les *blocks* de classifications congruentes avec le stéréotype (dans notre exemple, $B3$ et $B4$) et les *blocks* de classifications incongruentes avec le stéréotype (dans notre exemple, $B6$ et $B7$), est utilisée pour le calcul final du score. Cependant, si le pourcentage des temps de réponse trop rapides (e.g., inférieurs à 300 ms) dans l'ensemble des phases d'entraînement et de test est supérieur à 10%, alors le score n'est pas calculé et est invalide. Cela indique en effet que l'individu n'a pas passé le test sérieusement, et a appuyé très rapidement sur les touches du clavier pour classer les mots le plus vite possible, sans faire attention aux erreurs de classification. Pour le calcul du score, chaque temps de réponse trop rapide (e.g., inférieur à 400 ms) ou trop lent (e.g., supérieur à 10000 ms) est écarté. Deux sous-scores, Da et Db, sont alors calculés. Da correspond au score du premier *block* incongruent ($B6$) et du premier *block* congruent ($B3$), alors que Db correspond au score du deuxième *block* incongruent ($B7$) et du deuxième *block* congruent ($B4$). Pour ces deux sous-scores, la moyenne des temps de réponses du *block* congruent est soustraite à la moyenne des temps de réponses du *block* incongruent. L'estimateur de l'écart-type des temps de réponse non-exclus du *block* congruent et du *block* incongruent est alors divisé à la différence obtenue précédemment. Le score final du TAI, $Dscore$, correspond à la moyenne de ces deux sous-scores.

Nous avons prétesté notre TAI auprès de quatre élèves de 6e (non-REP) de notre entourage, et nous avons demandé l'avis des enseignants de mathématiques d'un collège en REP+. Dans l'ensemble, les quatre élèves n'avaient pas compris l'objectif du test et l'ont trouvé trop long, notamment à cause du côté répétitif du test. Un élève n'avait pas remarqué que les catégories

Algorithm 1 Algorithme calculant le score $Dscore$ du TAI.

procedure SCORE TAI($B1, B2, B3, B4, B5, B6, B7$)
  if (% de $TRs < 300$ ms $\forall$ *blocks*) > 10 % **then**
    score invalide. arrêt.
  end if
  for each $TR \in \{B3, B4, B6, B7\}$ **do**
    if $TR < 400$ ms or $TR > 10000$ ms **then**
      exclure TR
    end if
  end for
  $Da \leftarrow (avg(B6) - avg(B3))/\sigma(B3, B6)$
  $Db \leftarrow (avg(B7) - avg(B4))/\sigma(B7, B4)$
  $Dscore \leftarrow avg(Da, Db)$
  return $Dscore$
end procedure

Note. TR : Temps de Réponse. avg : moyenne. $B3$: premier *block* congruent. $B4$: deuxième *block* congruent. $B6$: premier *block* incongruent. $B7$: deuxième *block* incongruent. Les autres *blocks* ($B1, B2, B5$) sont des *blocks* d'entraînement. σ : estimateur de l'écart-type des temps de réponse non-exclus des deux *blocks* en entrée de la fonction.

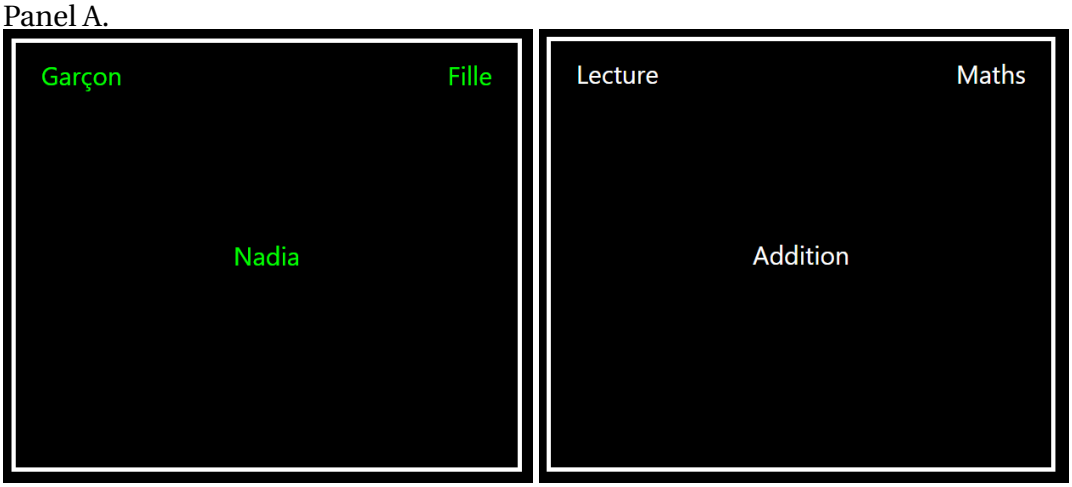
avaient changé lors du passage à l'étape suivante. Les élèves ont tout de même indiqué qu'ils avaient compris les instructions. Cependant, les enseignants nous ont fait remarquer que le test était difficile à comprendre pour des élèves de 6e et de 5e. Nous avons donc rajouté des exemples lors des instructions (voir Annexe K), et nous avons donné des explications à l'oral avant que les élèves commencent l'étude. Nous n'avons pas pré-testé ces nouvelles instructions. Lors de l'étude, l'expérimentatrice a expliqué que des mots allaient apparaître au centre de l'écran, et qu'il fallait les classer dans des catégories situées en haut de l'écran à l'aide des touches E et I du clavier (ou des flèches sur l'écran tactile de la tablette), le plus rapidement possible tout en faisant le moins d'erreur possible. L'expérimentatrice a énoncé les catégories, et a précisé qu'il fallait classer des prénoms pour les catégories "fille" et "garçon", et des mots simples comme "addition" ou "livre" pour les catégories "maths" et "lecture". Les élèves ont été prévenus que les mots apparaîtront plusieurs fois, et qu'il y avait plusieurs étapes dans cette tâche. L'expérimentatrice a précisé qu'il ne s'agissait pas d'un test de français, et qu'une voix prononcerait les mots à mesure qu'ils apparaîtront à l'écran pour ne pas que les élèves passent beaucoup de temps à lire. En effet, il était important de désamorcer l'idée qu'ils passeraient un test de français avant un test de mathématiques, car des travaux ont montré que cet ordre de passation réduisait les effets de menace du stéréotype en mathématiques chez les filles (SMEDING et al., 2013). Dû à la rapidité exigée pour le TAI, il a été précisé aux élèves qu'il était normal de faire quelques erreurs dans la précipitation. Si c'était le cas, une croix rouge apparaîtrait au-dessus du mot, et il faudrait alors classer le mot correctement. L'expérimentatrice a donné 4 exemples au tableau de la salle, illustrés en Figure 9.4 : 2 exemples lorsqu'il y a 2 catégories (phase d'entraînement), et 2 autres exemples lorsque les 4 catégories sont représentées simultanément (phase de test). Il a été précisé que le TAI n'était pas noté, mais que s'il était fait sérieusement, cela ajouterait un point à leur note au test de mathématiques. L'expérimentatrice a bien insisté sur l'importance d'aller le plus rapidement possible tout

en faisant le moins d'erreurs possible. Dans le cas contraire (faire trop d'erreurs, aller trop lentement ou rapidement), le score à ce test ne serait pas pris en compte et le bonus ne serait pas accordé. Elle a conseillé aux élèves de garder leurs index au-dessus de chacune de ces touches du clavier pour aller le plus rapidement possible.

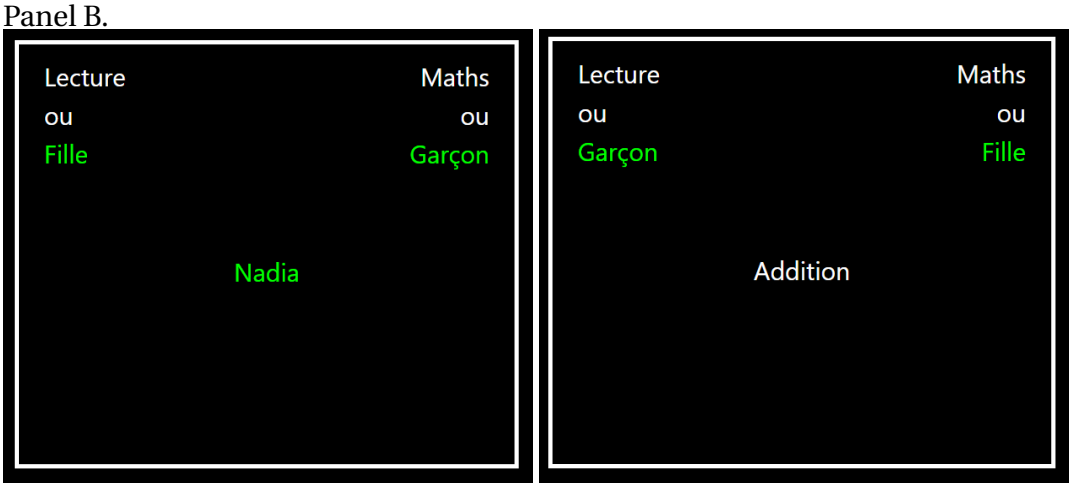
FIGURE 9.4.

Exemples du TAI genre-mathématiques-lecture donnés au tableau de la salle, lorsqu'il n'y a que deux catégories (panel A) et lorsque les quatre catégories sont présentes (panel B).

Panel A.



Panel B.



Notre objectif initial était d'utiliser les deux scores obtenus à ce TAI aux temps 1 et 2 pour tester la consistance interne de ce nouvel outil en test-retest, puis d'utiliser la moyenne des deux scores comme modérateur potentiel des effets des agents sur la performance et les auto-évaluations en mathématiques au temps 2. Or, plusieurs résultats nous ont contraint à réviser cette stratégie. Tout d'abord, lors de la première passation du TAI (temps 1), malgré les explications et exemples donnés au tableau, les élèves n'ont pas bien compris ce qu'ils devaient faire et ont interrompu la tâche à plusieurs reprises pour poser des questions. En

revanche, lors de la deuxième passation quelques semaines plus tard (temps 2), les élèves qui étaient présents lors de la première passation ont passé le TAI sans problème particulier. En effet, le fait d'avoir déjà passé le TAI une première fois et d'avoir eu l'opportunité d'obtenir de plus amples explications sur la tâche, a conduit à un pourcentage plus faible de scores invalides au temps 2 (7.5%) par rapport aux élèves qui ont découvert le TAI pour la première fois au temps 2 (11.4%). Ensuite, la corrélation entre les deux sous-scores du TAI, Da et Db, est meilleure au temps 2, $r(326)=.79$, $p<.001$, qu'au temps 1, $r(399)=.65$, $p<.001$. Enfin, la corrélation test-retest entre les deux $Dscores$ du TAI est non significative (bien que positive), $r(326)=.07$, $p=.180$. Pris ensemble, ces résultats invitent à ne pas s'attarder sur le score du TAI obtenu lors de la première passation, et à plutôt exploiter le score du TAI obtenu au temps 2, en ne retenant bien sûr que les élèves ayant obtenu un score valide et l'ayant déjà passé au temps 1.

9.2.3.3. Performances aux tests standardisés de mathématiques

Tous les tests standardisés de mathématiques proposés ont été créés en collaboration avec les enseignants en mathématiques d'un collège en REP+. Les 6e et les 5e ont chacun des tests différents, portant sur des notions différentes (le parallélisme en 6e, et les priorités opératoires en 5e - voir Chapitre 8 pour les détails des tests). Les tests standardisés de mathématiques proposés lors des deux temps portaient sur les mêmes notions, mais différaient en termes de nombre d'exercices, de formulation des questions, et de points alloués. Par conséquent, pour chacun des deux tests, 3 scores ont été exploités :

- Le score globale d'exactitude.
- Deux sous-scores :
  - Un score d'exactitude pour les exercices faciles (taux de réussite ≥ 50%).
  - Un score d'exactitude pour les exercices difficiles (taux de réussite < 50%) [1].

9.2.3.4. Moyenne trimestrielle en mathématiques

Nous avons recueilli auprès des établissements scolaires la moyenne des élèves en mathématiques au premier trimestre afin d'utiliser cette variable en covarié dans nos analyses principales. Or, une différence de pratique dans le type de notation selon les établissements de notre échantillon s'est avérée intéressante et importante à considérer. En effet, l'évaluation la plus courante dans les collèges en France est l'attribution d'une note allant de 0 à 20 pour chaque contrôle, où les enseignants calculent ensuite la moyenne de toutes ces notes pour attribuer un score dans le domaine. Nous avons appelé ce type d'évaluation la notation classique. Cependant, certains collèges n'attribuent plus de note et affectent à la place une couleur pour indiquer le groupe de compétence auquel l'élève appartient. Au sein de notre échantillon, 16 classes utilisent la notation dite classique, et 7 classes en REP+ utilisent le type de notation par couleur (soit 58.33% des classes de REP+ et 0% des classes non-REP). Cette dernière distingue trois couleurs/catégories : *insuffisant* (rouge) correspondant à une moyenne strictement inférieure à 8, *moyen* (orange) pour une moyenne comprise entre 8 et 15 inclus, et *bon* (vert) avec une moyenne strictement supérieure à 15. Compte tenu du nombre

1. Les exercices *t1_exo2_2_e* (Annexe H) et *t2_exo2_2_e* (Annexe I) ont été retirés pour les tests de 5e, car ils se sont avérés trop difficiles pour tous les élèves (moins de 5% de réussite). Ils tiraient artificiellement le score vers le bas pour les 5e comparativement aux 6e, ce qui introduisait une variable confondue.

important de classes dans notre échantillon à utiliser le mode de notation par couleur, il nous a semblé important de tester, dans nos analyses préliminaires, si le type de notation (classique *versus* par couleur) avait une influence sur nos différentes variables (i.e., auto-évaluations en mathématiques, score TAI, moyennes trimestrielles et performances aux tests standardisés de mathématiques). Enfin, la conversion des trois couleurs en notes sur 20 a été réalisée avec l'assistance des enseignants. Le niveau insuffisant (rouge) correspondrait à une moyenne de 8/20, tandis que le niveau moyen (orange) serait associé à une moyenne de 11.5/20, et le niveau bon (vert) à une moyenne de 15/20. Bien que non optimale, cette conversion était le seul moyen de tenir compte des niveaux préalables en mathématiques des élèves dans nos analyses principales.

9.3. Résultats

9.3.1. Analyses préliminaires

Les objectifs de ces analyses sont de présenter les statistiques descriptives des variables de l'étude (auto-évaluations, score TAI, performances aux tests standardisés), d'analyser les effets d'une variable non anticipée dans l'étude (i.e., le type de notation des enseignants : classique *versus* par couleur), et de vérifier qu'il n'y ait aucun biais de répartition des élèves dans les conditions expérimentales au temps 1 puis au temps 2 en fonction de leurs niveaux préalables en mathématiques.

9.3.1.1. Statistiques descriptives

Le Tableau 9.3 présente les moyennes, erreurs standards, et corrélations bivariées entre l'ensemble des mesures au temps 1 et au temps 2. La variable du sexe des participant corrélait négativement avec les deux scores de sentiment de réussite (temps 1 et temps 2), car les garçons ont été codés en 0 et les filles en 1. Les deux scores d'auto-évaluation, à savoir le sentiment de réussite en mathématiques et l'importance accordée aux mathématiques, corrélaient positivement entre eux : plus les élèves trouvaient les mathématiques importantes, plus ils pensaient réussir en mathématiques, et vice-versa. Le sentiment de réussite en mathématiques corrélait positivement avec tous les scores de performance. L'importance accordée aux mathématiques corrélait avec les scores de performance au temps 1, et l'importance mesurée au temps 2 corrélait positivement avec les scores de performance aux temps 1 et 2, toujours à l'exception de la performance sur les exercices faciles. Plus les élèves considéraient les mathématiques comme importantes, mieux ils réussissaient dans ce domaine (et vice-versa).

TABLEAU 9.3

Scores moyens, erreurs standards, et corrélations bivariées entre l'ensemble des mesures.

Variables	M	ES	1	2	3	4	5	6	7	8	9	10	11	12
1. Sexe de l'élève	.	.	-											
2. Sentiment de réussite en mathématiques, T1	3.84	.04	-.25**	-										
3. Sentiment de réussite en mathématiques, T2	3.78	.04	-.31**	.75**	-									
4. Importance accordée aux mathématiques, T1	4.40	.03	-.01	.35**	.23**	-								
5. Importance accordée aux mathématiques, T2	4.34	.04	.02	.32**	.32**	.61**	-							
6. Performance (globale), T1	42.72	1.05	-.06	.28**	.28**	.12*	.13*	-						
7. Performance (facile), T1	69.45	1.42	-.06	.14*	.12*	.10	.10	.61**	-					
8. Performance (difficile), T1	31.70	1.24	-.02	.28**	.26**	.15**	.15**	.92**	.37**	-				
9. Performance (globale), T2	46.52	1.15	-.07	.27**	.27**	.06	.15**	.59**	.51**	.56**	-			
10. Performance (facile), T2	71.62	1.29	-.05	.20**	.21**	.05	.09	.44**	.48**	.38**	.81**	-		
11. Performance (difficile), T2	29.29	1.34	-.06	.25**	.26**	.07	.16**	.56**	.42**	.57**	.92**	.51**	-	
12. TAI, T2	.12	.03	-.10	-.05	-.04	-.04	-.10	.08	.11*	.06	.09	.11	-05	-

Note. T1=Temps 1, T2=Temps 2. La variable catégorielle a été codée de la manière suivante : Sexe de l'élève : 0=garçon, 1=fille.
*p<.05. **p<.01.

9.3.1.2. Analyses sur les effets potentiels du type de notation en classe

Comme nous l'avons énoncé, une variable non-anticipée dans l'étude nous a poussé à explorer si elle constituait un biais pour les analyses suivantes. Il s'agit du type de notation utilisée par les enseignants pour évaluer leurs élèves lors des contrôles en classe. Pour les analyses subséquentes, nous avons codé cette variable en nominale, avec deux modalités : notation classique *versus* notation par couleur. Nous examinons ici si le type de notation influençait les moyennes trimestrielles en mathématiques, les auto-évaluations en mathématiques, la stéréotypie implicite, et les performances aux tests standardisés de mathématiques. La répartition, même très déséquilibrée, des filles et des garçons en fonction du type de notation (Tableau 9.4) a permis de réaliser des ANOVAs intersujets 2 (type de notation : classique/couleur) X 2 (sexe de l'élève : fille/garçon) pour tester ces effets potentiels.

TABLEAU 9.4

Nombre de participants selon le type de notation, aux temps 1 et 2.

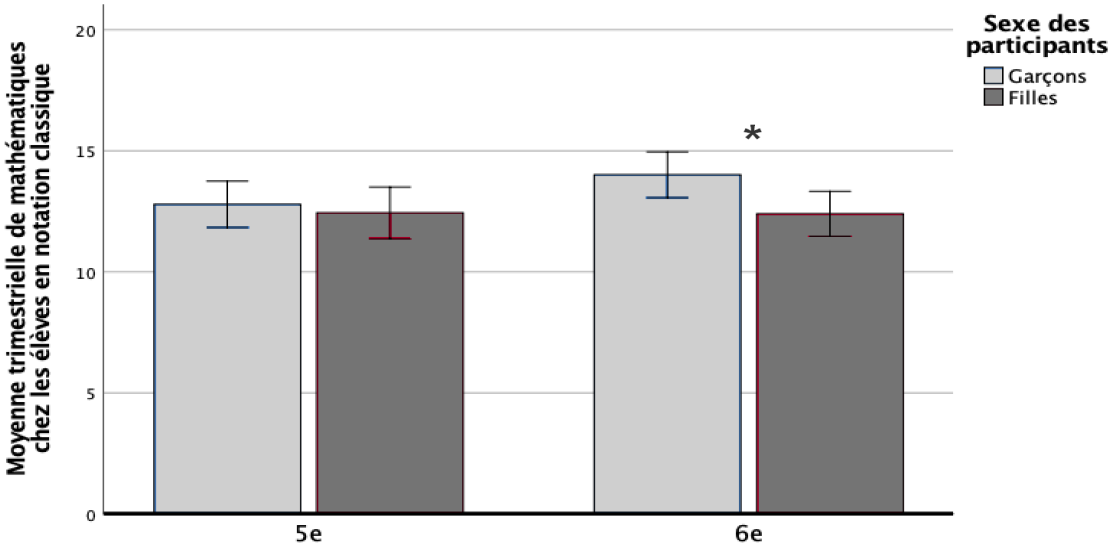
Sexe des participants	Type de notation	
	Classique	Couleur
Temps 1		
Garçon	152	60
Fille	141	58
Temps 2		
Garçon	142	55
Fille	124	54

Moyennes trimestrielles en mathématiques. Nous avons tout d'abord réalisé l'ANOVA 2 (type de notation) X 2 (sexe de l'élève) sur l'échantillon complet englobant à la fois les REP+ et les non-REP. Les résultats ont indiqué uniquement un effet principal du type de notation, $F(1, 404)=9.90$, $p=.002$, $\eta^2_p=.02$. Pour les filles et les garçons, la notation classique ($M=12.91$, $ES=.22$) se traduisait par de meilleures moyennes en mathématiques que la notation par couleur ($M=11.64$, $ES=.34$). La même analyse a été conduite uniquement sur les établissements REP+ pour voir si l'influence du type de notation demeurait en enlevant les collèges non-REP qui regroupent généralement des élèves dont le niveau scolaire est un peu meilleur. L'effet principal n'était plus significatif, $F(1, 309)=1.68$, $p=.196$, $\eta^2_p=.01$, mais les moyennes allaient dans le même sens (classique : $M=12.18$, $ES=.26$; couleur : $M=11.64$, $ES=.33$). Nous avons ensuite séparé les élèves notés de manière classique et les élèves notés avec une couleur pour déterminer s'il y avait des différences de moyenne trimestrielle entre les filles et les garçons. L'ANOVA 2 (classe) X 2 (sexe de l'élève) sur l'échantillon des élèves notés de manière classique a révélé un effet principal du sexe de l'élève, $F(1, 286)=3.90$, $p=.049$, $\eta^2_p=.01$. Les garçons ($M=13.40$, $ES=.34$) avaient de meilleures notes en cours que les filles ($M=12.42$, $ES=.36$). Cette différence était présente en 6e (garçons : $M=14.01$, $ES=.48$; filles : $M=12.40$, $ES=.47$), $F(1, 286)=5.72$, $p=.017$, $\eta^2_p=.02$, mais pas en 5e, (garçons : $M=12.79$, $ES=.49$; filles : $M=12.44$, $ES=.54$), $F(1, 286)=.23$, $p=.630$, $\eta^2_p<.01$, à cause d'une baisse de performance des garçons (Figure 9.5). Pour les élèves notés avec des couleurs, le test du χ^2 entre les groupes de niveaux et le sexe des élèves n'a pas montré de lien entre les deux variables, $\chi^2(2, 118)=2.04$, $p=.361$, $V=.13$, bien que plus de garçons étaient présents parmi les très bons élèves (65.00%)

et plus de filles étaient représentées parmi les élèves ayant des résultats insuffisants (55.60%). Il faut cependant noter que la très grande majorité des élèves se trouvaient dans la catégorie "bon", ce qui n'est pas étonnant puisqu'elle regroupe les élèves qui se situeraient entre 8 et 15 sur 20. La Figure 9.6 montre les effectifs des filles et des garçons selon les trois groupes de niveaux. Une absence de différence entre filles et garçons dans les moyennes trimestrielles en mathématiques ne signifie cependant pas une absence de menace du stéréotype, puisque les filles peuvent compenser la réduction de leur capacité de mémoire de travail lors des tests non-standardisés.

FIGURE 9.5.

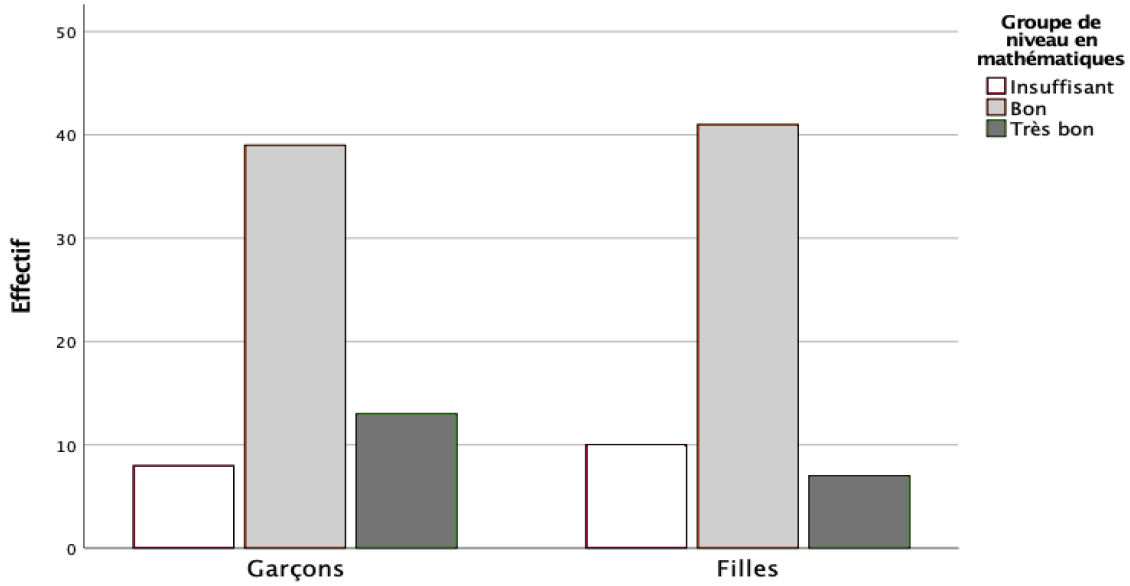
Moyennes trimestrielles en mathématiques des filles et des garçons selon leurs classes, notées de manière classique (note sur 20).



Note. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

FIGURE 9.6.

Effectif des filles et des garçons dans les trois groupes de niveaux de la notation en couleur.



Performances aux tests standardisés de mathématiques. Concernant les performances aux tests standardisés de mathématiques aux temps 1 et 2, les ANOVAs ont révélé un effet principal du type de notation sur tous les scores, qui indiquait systématiquement de meilleures performances avec la notation classique (cf Tableau 9.5), à l'exception des performances aux exercices difficiles au temps 1. Probablement que ces exercices étaient difficiles pour tout le monde, bien que l'on a encore observé une moins bonne performance de la part des élèves notés en couleur (non-significative).

TABLEAU 9.5

Scores moyens et erreurs standards pour chaque score de performance, au temps 1 et au temps 2, selon le type de notation en classe (classique ou couleur).

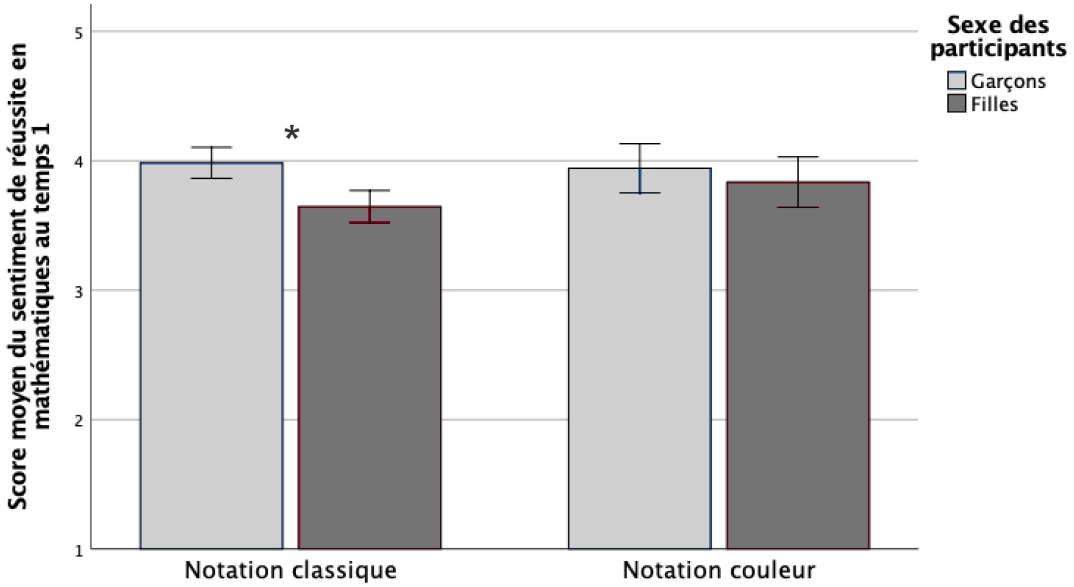
Performances	Classique		Couleur		F	p	η_p^2
	M	ES	M	ES			
temps 1							
Globale	44.56	1.17	38.12	1.83	8.79	.003	.02
Facile	73.06	1.57	60.47	2.48	18.40	<.001	.04
Difficile	32.96	1.38	28.59	2.18	2.87	.091	.01
temps 2							
Globale	49.12	1.35	39.93	2.11	13.52	<.001	.04
Facile	74.95	1.52	63.38	2.37	16.93	<.001	.04
Difficile	31.42	1.59	23.83	2.48	6.62	.010	.02

Sentiment de réussite en mathématiques. Lors du temps 1, les analyses ont montré uniquement un effet principal du sexe des élèves, $F(1, 407)=7.31$, $p=.007$, $\eta_p^2=.02$. Les filles ($M=3.74$, $ES=.06$) s'auto-évaluaient moins positivement que les garçons en mathématiques ($M=3.96$, $ES=.06$), ce qui est cohérent avec les résultats habituels en sciences sociales. Nous avons observé le même effet principal au temps 2, $F(1, 370)=17.24$, $p<.001$, $\eta_p^2=.05$, avec le même résultat (filles : $M=3.59$, $ES=.06$; garçons : $M=3.95$, $ES=.06$). Même si l'interaction type de notation X sexe de l'élève n'était pas significative au temps 1, il est intéressant de noter que les auto-évaluations moins positives des filles ($M=3.65$, $ES=.06$) comparées aux garçons ($M=3.98$, $ES=.06$) s'observaient surtout dans le cas de la notation classique, le contraste étant significatif, $F(1, 407)=14.79$, $p<.001$, $\eta_p^2=.05$. La notation par couleur semble lisser les différences filles/garçons. Cette tendance était confortée par l'interaction au temps 2, cette fois-ci significative, $F(1, 370)=5.45$, $p=.020$, $\eta_p^2=.02$. On retrouve le contraste significatif entre filles ($M=3.47$, $ES=.07$) et garçons ($M=4.04$, $ES=.07$) dans la condition de notation classique, $F(1, 370)=36.01$, $p<.001$, $\eta_p^2=.09$, et une tendance de la notation par couleur ($M=3.70$, $ES=.11$) à booster les auto-évaluations des filles par rapport à la notation classique ($M=3.47$, $ES=.07$), $F(1, 370)=3.39$, $p=.066$, $\eta_p^2=.01$. Les graphiques en Figure 9.7 illustrent l'interaction entre les deux variables indépendantes, au temps 1 et au temps 2.

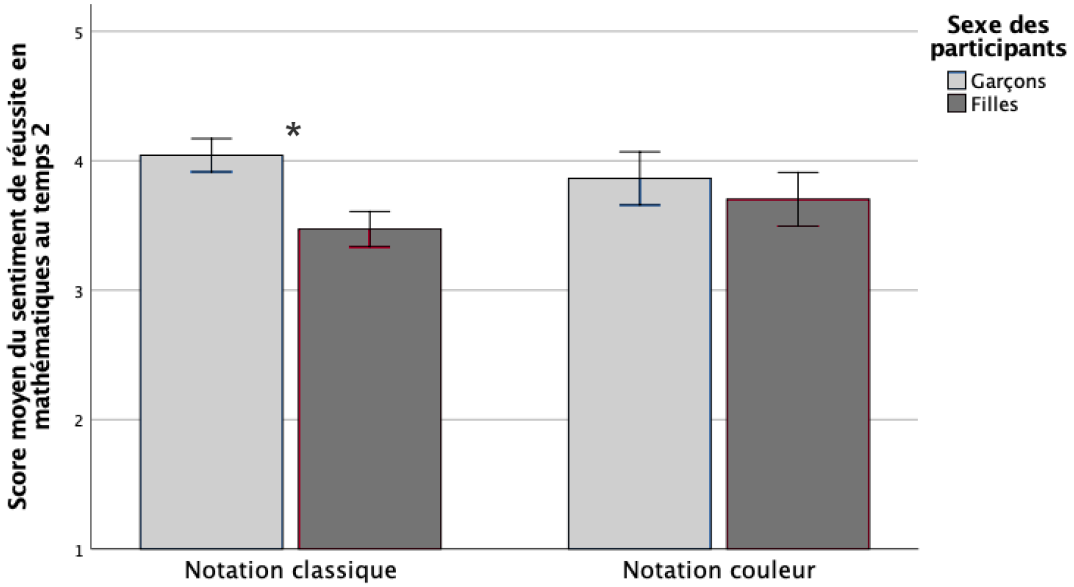
FIGURE 9.7.

Score moyen du sentiment de réussite en mathématiques chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves, au temps 1 (panel A) et au temps 2 (panel B).

Panel A.



Panel B.



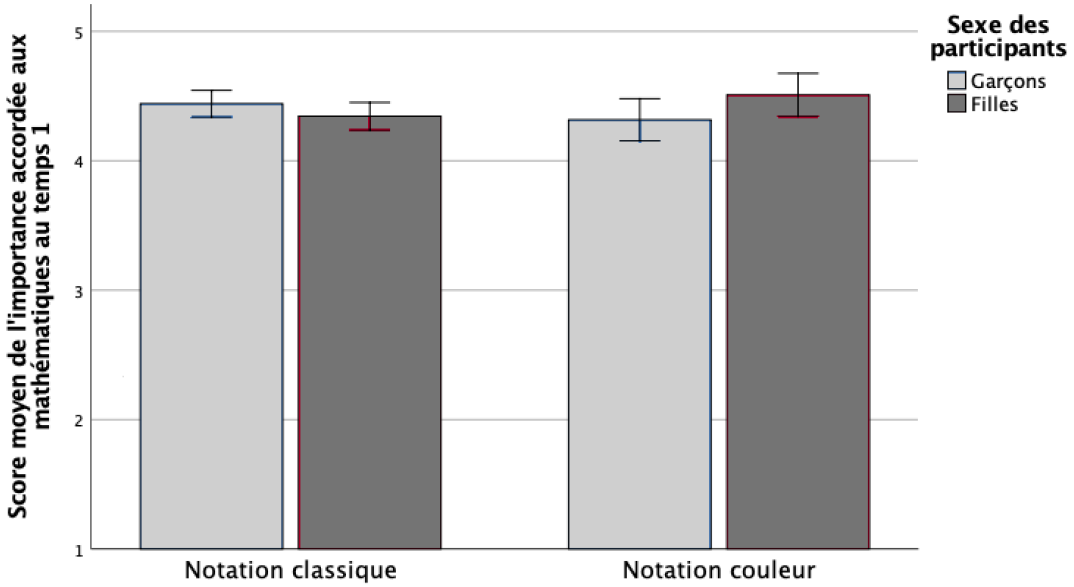
Note. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

Importance accordée aux mathématiques. Comme le montre la Figure 9.8, le type de notation n'impactait par cette évaluation qui était très élevée pour tous les élèves, au temps 1 comme au temps 2.

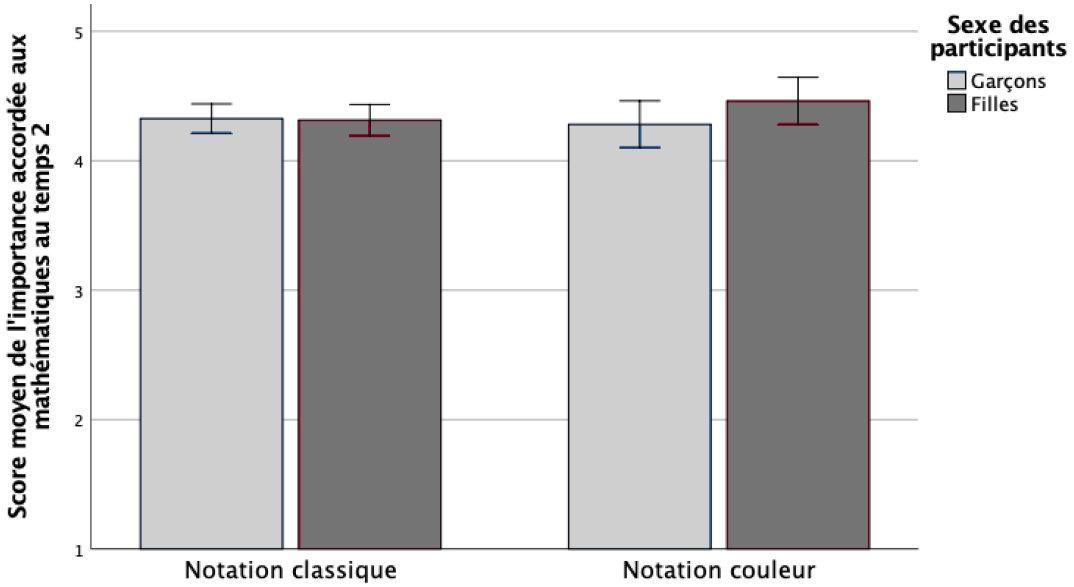
FIGURE 9.8.

Score moyen de l'importance accordée aux mathématiques chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves, au temps 1 (panel A) et au temps 2 (panel B).

Panel A.



Panel B.



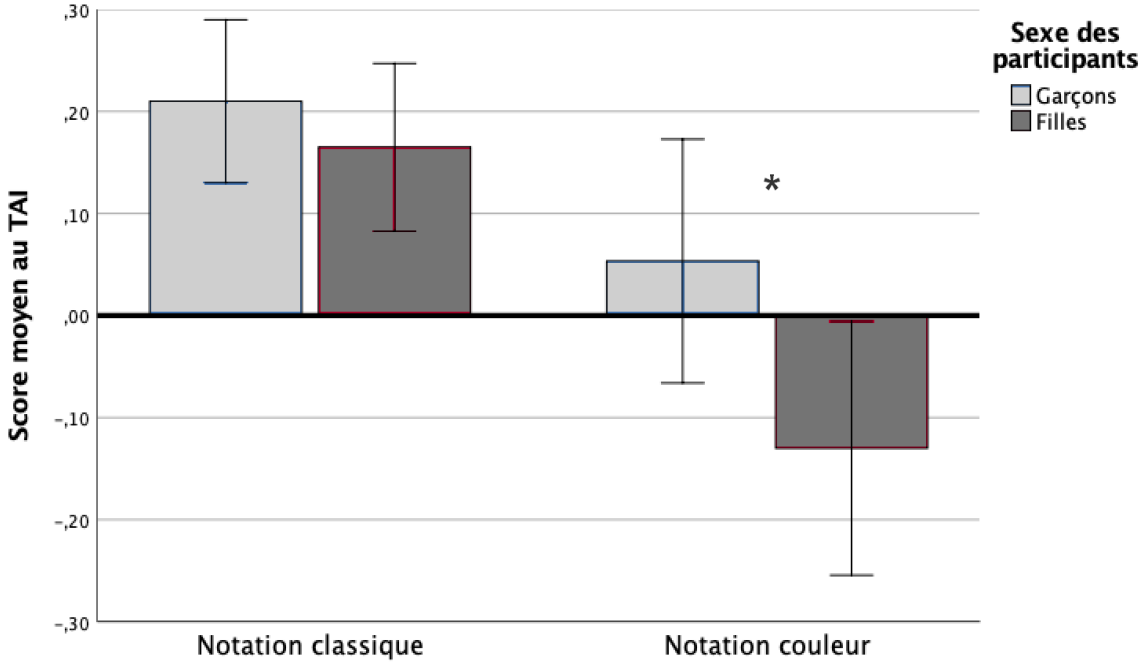
Note. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

Score de stéréotypie implicite. L'ANOVA 2 (type de notation) X 2 (Sexe de l'élève) a été réalisée uniquement sur les scores TAI obtenus au temps 2, et uniquement sur les élèves ayant obtenu un score valide aux deux temps de mesure afin de garantir une compréhension et passation correctes du test. Les résultats ont indiqué que les deux effets principaux étaient

significatifs (sexe de l'élève : $F(1, 322)=18.42$, $p<.001$, $\eta^2_p=.05$; type de notation : $F(1, 322)=4.71$, $p=.031$, $\eta^2_p=.01$). Le score du TAI était positif et plus élevé en condition de notation classique ($M=.19$, $ES=.03$) qu'en condition d'évaluation par couleur ($M=-.04$, $ES=.04$) où il était proche de zéro. Bien que peu élevé, le score positif de .19 montre que les élèves en notation classique avaient des automatismes stéréotypiques, comparés aux élèves en notation par couleur qui n'en avaient pas (ni stéréotypique, ni contre-stéréotypique). Quant à l'effet principal du sexe des élèves, il indique que les associations automatiques entre garçons et mathématiques et entre filles et lecture étaient plus fortes chez les garçons ($M=.13$, $ES=.04$) que les filles ($M=.02$, $ES=.04$). Même si l'interaction n'était pas significative, il est intéressant de noter la similitude des scores TAI (et positifs) chez les filles et garçons en notation classique, alors que la notation par couleur exacerbe l'écart de score entre les deux groupes, avec un score traduisant une contre-stéréotypie chez ces filles (Figure 9.9).

FIGURE 9.9.

Score moyen au TAI genre-mathématiques-lecture au temps 2 chez les filles et les garçons, en fonction du type de notation utilisé par les enseignants pour évaluer les élèves.



Note. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

Discussion. Pris ensembles, les résultats ci-dessus sont assez cohérents et intéressants. Le type de notation induirait des différences dans les auto-évaluations de compétences, le niveau de stéréotypie implicite et les performances aux tests standardisés de mathématiques. Comparativement à la notation classique, la notation par couleur semble booster les auto-évaluations de compétences en mathématiques des filles et neutraliserait ainsi les différences habituellement observées entre filles et garçons dans cette matière. La notation par couleur était également associée à un niveau plus faible de stéréotypie implicite dans le domaine, chez tous les élèves et en particulier chez les filles. Un mécanisme explicatif probable repose

sur le processus de comparaisons sociales évaluatives intergroupes qui est très présent dans le cas de la notation classique, très compétitive. Les évaluations comparatives sont au cœur des différences filles/garçons et de la construction des représentations stéréotypiques. La notation par couleur permettrait peut-être d'échapper à ces effets évaluatifs comparatifs et hiérarchiques, où ce processus de comparaisons serait moins présent. En contrepartie, les performances sont moins bonnes, probablement en raison du processus comparatif qui serait moins présent, indispensable pour situer son niveau de compétence et générer aussi de la motivation compétitive à la performance.

Ces résultats nous ont incité à prendre en compte le type de notation lors de nos analyses centrales. Nous avons fait le choix de mettre le type de notation en covarié dans toutes les analyses. Il était en effet impossible d'utiliser le type de notation en variable indépendante au temps 2, en raison d'un nombre insuffisant (voir nul) de participants dans certaines conditions expérimentales.

9.3.1.3. Analyses sur les biais de répartition des élèves dans les conditions expérimentales en fonction de leur niveau initial en mathématiques

Ces analyses ont pour but de vérifier qu'il n'y a pas eu de biais de répartition des élèves dans les conditions expérimentales aux temps 1 et 2 en fonction de leur niveau initial en mathématiques (moyenne trimestrielle). Les moyennes en mathématiques des élèves obtenues dans leur classe au premier trimestre devraient en effet être indépendantes des conditions expérimentales du temps 1 (voix féminine *versus* masculine de la leçon) et du temps 2 (genre et type de comportement de l'agent virtuel). Un éventuel effet significatif des conditions expérimentales traduirait une répartition biaisée dont il faudrait tenir compte d'un point de vue statistique.

Au temps 1, l'ANOVA intersujets 2 (sexe de l'élève : fille/garçon) X 2 (voix de la leçon : masculine/féminine) réalisée sur la moyenne trimestrielle en mathématiques des élèves (avec le niveau 6e/5e et le type de notation en covariés) a révélé un effet principal de la voix de la leçon, $F(1, 402)=4.69$, $p=.031$, $\eta^2_p=.01$: les élèves en condition de voix masculine ($M=12.95$, $ES=.26$) avaient une moyenne trimestrielle en mathématiques un peu plus élevée que ceux en condition de voix féminine ($M=12.14$, $ES=.26$). Cette répartition biaisée des élèves dans les deux conditions expérimentales invite à mettre la moyenne trimestrielle en mathématiques comme covarié dans les analyses du temps 1.

Au temps 2, l'ANOVA intersujets 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : masculine/féminine) X 3 (type de l'agent : $MER/M\overline{ER}/\overline{MER}$) réalisée sur la moyenne trimestrielle en mathématiques des élèves (avec le niveau 6e/5e et le type de notation en covariés) a révélé plusieurs effets d'interactions significatifs. Tout d'abord, l'effet de l'interaction entre le sexe de l'élève et le type de l'agent était significatif, $F(2, 358)=3.07$, $p=.047$, $\eta^2_p=.02$, ainsi que l'interaction entre le genre de l'agent et le type de l'agent, $F(2, 358)=3.24$, $p=.041$, $\eta^2_p=.02$. Enfin, un effet d'interaction triple significatif entre le genre de l'agent, le sexe de l'élève, et le type de l'agent a été révélé, $F(2, 358)=3.72$, $p=.025$, $\eta^2_p=.02$. Pour contrôler ces biais de répartition, nous allons utiliser en covariés : la moyenne trimestrielle en mathématiques, son interaction avec le genre de l'agent, son interaction avec le type de l'agent, et son interaction avec le genre

et le type de l'agent.

9.3.2. Analyses centrales

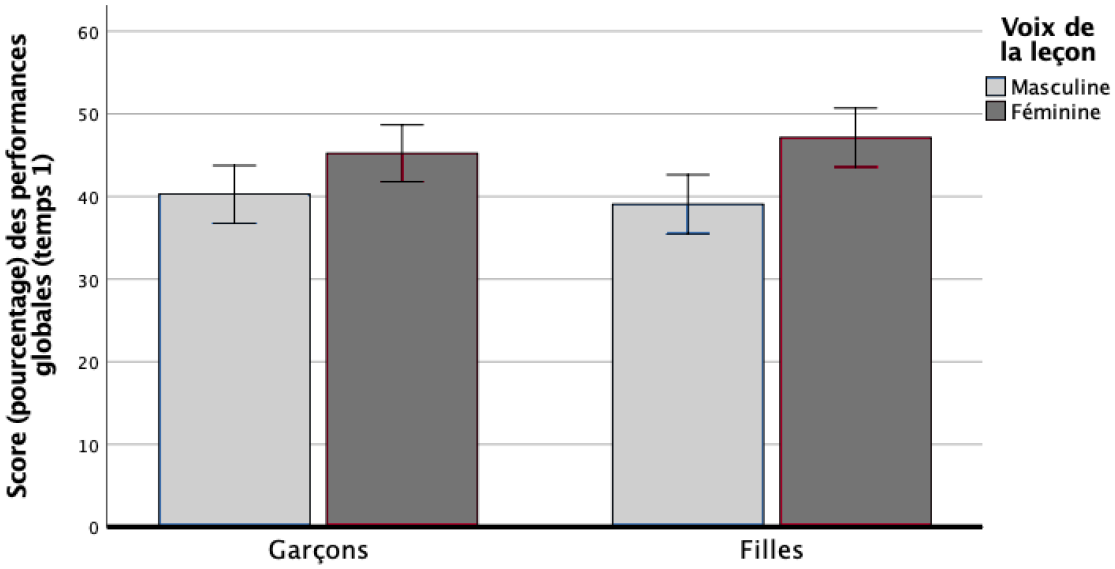
9.3.2.1. Analyses au temps 1

Pour rappel, le temps 1 avait pour objectif de familiariser les élèves avec la passation des tests de mathématiques sur ordinateur ou tablette et d'obtenir une *baseline* pour la stéréotypie implicite. Les élèves n'étaient pas exposés aux agents, mais la voix féminine *versus* masculine des agents était utilisée pour lire la leçon et donner les consignes avant le test de mathématiques. Bien que nous n'avions pas d'attente particulière sur cette première phase, nous avons exploré si le genre de la voix qui délivre la leçon avait influencé les performances des élèves au test standardisé de mathématiques.

Une ANOVA intersujets 2 (sexe de l'élève : fille/garçon) X 2 (voix de la leçon : masculine/féminine) avec la moyenne trimestrielle en mathématiques, la classe (6e/5e) et le type de notation en covariés, a tout d'abord été réalisée sur le score global obtenu au test de mathématiques. L'analyse a révélé uniquement un effet principal de la voix de la leçon, $F(1, 401)=12.83$, $p<.001$, $\eta^2_p=.03$. Comparativement à la voix masculine ($M=39.66$, $ES=1.27$), la voix féminine ($M=46.18$, $ES=1.27$) était associée à de meilleures performances. Il est intéressant de noter que cela va à l'encontre du biais de répartition initiale évoqué précédemment, où les élèves avec une meilleure moyenne trimestrielle en mathématiques se situaient dans la condition de la voix masculine. La Figure 9.10 illustre les performances des filles et des garçons, selon la voix de la leçon.

FIGURE 9.10.

Score moyen des performances globales chez les filles et les garçons au temps 1, en fonction de la voix de la leçon.



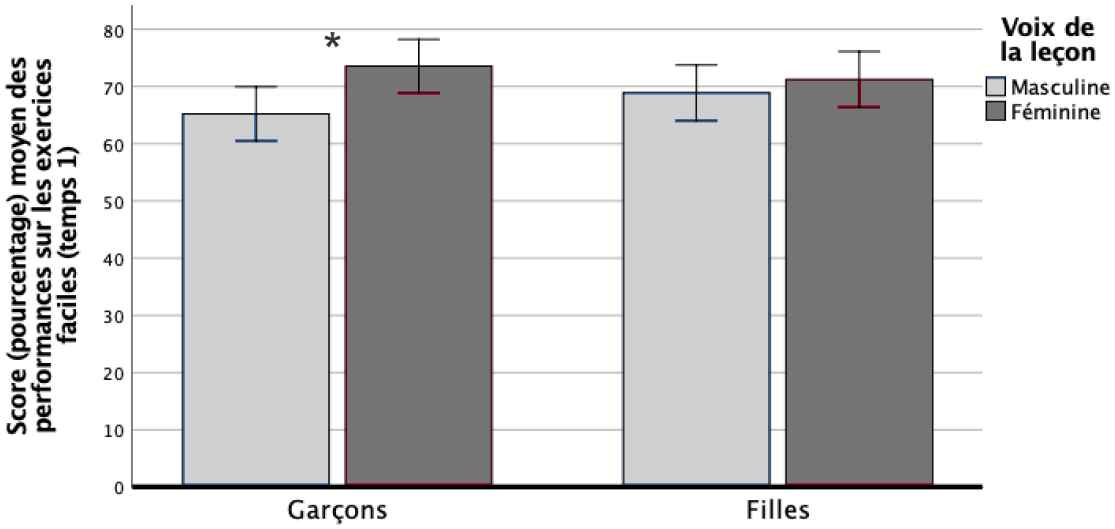
Note. La voix féminine est associée à de meilleures performances pour tous les élèves. Barres d'erreur : intervalle de confiance à 95%.

Nous avons réalisé une autre analyse pour tenir compte du niveau de difficulté des exercices (faciles *versus* difficiles). Une ANOVA mixte 2 (sexe de l'élève : fille/garçon) X 2 (voix de la leçon : masculine/féminine) X 2 (difficulté des exercices : facile/difficile), avec le dernier facteur en intrasujets, et les trois mêmes covariés précédents (moyenne trimestrielle en mathématiques, classe, type de notation), a été réalisée. Parmi les effets des facteurs intersujets, on retrouve l'effet principal de la voix de la leçon, où la voix féminine était plus bénéfique pour les élèves, $F(1, 401)=11.48$, $p<.001$, $\eta^2_p=.03$. Les analyses ont également montré un effet d'interaction entre le sexe de l'élève, la voix de la leçon, et le niveau de difficulté, $F(1, 401)=3.98$, $p=.047$, $\eta^2_p=.01$ (Figure 9.11). L'effet bénéfique de la voix féminine ($M=36.75$, $ES=2.16$) comparativement à la voix masculine ($M=27.14$, $ES=2.16$) s'observait chez les filles pour les exercices difficiles, $F(1, 401)=9.75$, $p=.002$, $\eta^2_p=.02$, mais aussi chez les garçons pour les exercices faciles (voix féminine : $M=73.57$, $ES=2.38$; voix masculine : $M=65.21$, $ES=2.41$), $F(1, 401)=6.03$, $p=.014$, $\eta^2_p=.02$. La voix féminine semble réduire les effets de menace du stéréotype, en agissant directement sur les exercices difficiles où il est plus habituel d'observer une contre-performance de la part des filles. Cette voix a également un effet positif pour les garçons, mais uniquement pour les exercices faciles.

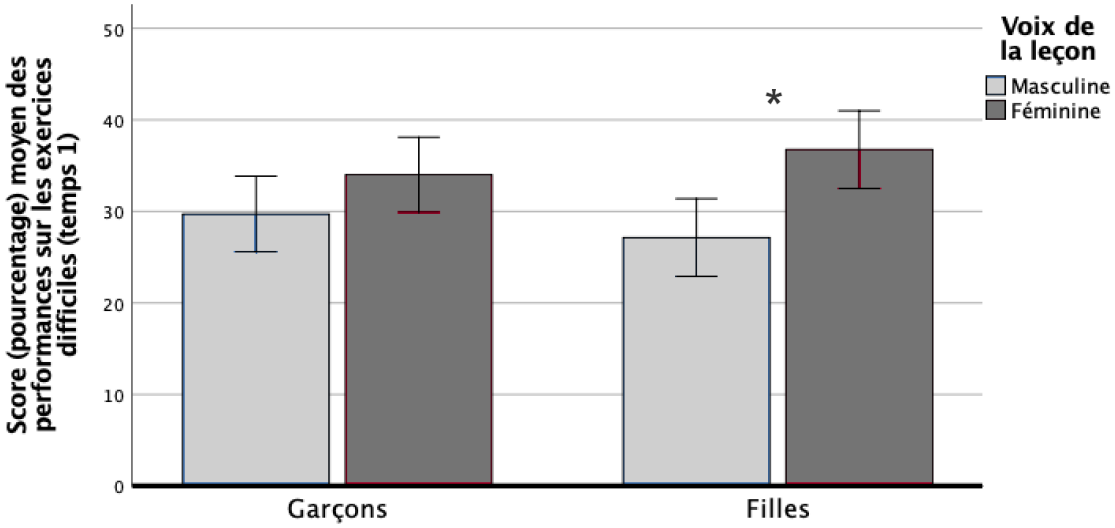
FIGURE 9.11.

Score moyen des performances chez les filles et les garçons au temps 1 sur les exercices faciles (panel A) et difficiles (panel B), en fonction de la voix de la leçon.

Panel A.



Panel B.



Note. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

9.3.2.2. Analyses au temps 2

Le temps 2 avait pour objectif d'introduire les agents virtuels pédagogiques, afin de mesurer leur impact sur les performances en mathématiques des élèves, et plus particulièrement des filles. Les élèves ont interagi avec un agent, féminin ou masculin, parmi l'un des trois types

TABLEAU 9.6

Nombre de participants inclus dans les analyses au temps 2, par conditions expérimentales.

Genre de l'agent	Sexe du participant	Type de l'agent		
		$M\overline{ER}$	$ME\overline{R}$	MER
Masculin	Garçons	30	22	29
	Filles	27	29	22
Féminin	Garçons	27	32	28
	Filles	25	28	27

d'agents que nous avons créés : l'agent avec comportements de rapport et justifiant sa réussite en mathématiques par ses efforts et son travail (MER), l'agent ayant ces mêmes explications quant à sa réussite en mathématiques mais sans comportements de rapport ($ME\overline{R}$), et l'agent n'ayant pas de comportements de rapport et ne donnant aucune explication quant à l'intelligence malléable ($M\overline{ER}$). Notre hypothèse était que l'agent avec des explications et des comportements de rapport virtuel serait associé à de meilleures performances en mathématiques, et de meilleures auto-évaluations en mathématiques. Nous n'avions pas d'hypothèse tranchée sur le genre de l'agent, puisque des études ont montré que ce sont uniquement des femmes ou des filles en tant que modèles qui ont pu réduire les effets de menace du stéréotype, alors que d'autres études ont montré que c'est l'explication de la réussite par le travail qui prime. Enfin, nous avons également analysé l'effet modérateur de la stéréotypie implicite sur l'impact des agents virtuels, sur les performances et les auto-évaluations en mathématiques. Toutes les analyses effectuées ont permis de tester nos hypothèses.

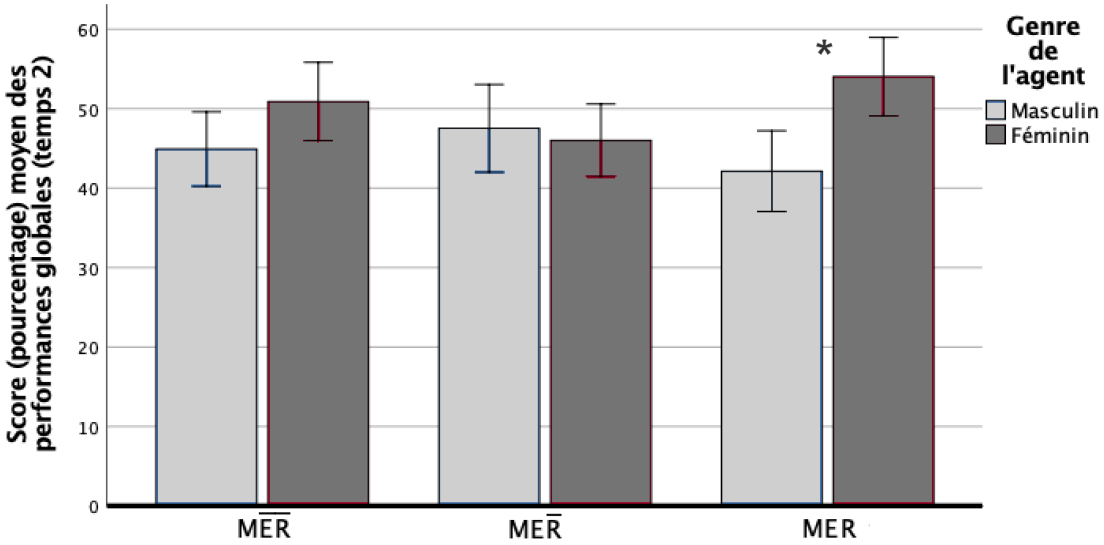
Pour ces analyses, nous avons inclus les élèves qui ont également participé au temps 1. Le premier temps permettait en effet aux élèves de s'habituer aux tâches que nous leur avons demandées (passations numériques du TAI et des tests de mathématiques). Nous avons également inclus uniquement les élèves ayant un score au TAI valide aux deux temps. Même si l'on a exploité le score au TAI uniquement au temps 2, le score valide aux deux temps de mesure est une précaution supplémentaire concernant le sérieux de ces élèves lors des passations. Le Tableau 9.6 présente la répartition des participants inclus dans nos analyses dans les différentes conditions expérimentales. Pour toutes les analyses de cette section, les covariés sont le type de notation (classique/couleur), la classe (6e/5e), et les 4 covariés liés aux moyennes trimestrielles en mathématiques (i.e., la moyenne trimestrielle en mathématiques, son interaction avec le genre de l'agent, son interaction avec le type de l'agent, et son interaction avec le genre et le type de l'agent) pour palier le biais de répartition dans les conditions expérimentales. Nous avons d'abord testé l'effet des agents sur les performances au test standardisé de mathématiques, puis sur les auto-évaluations, et enfin sur l'effet modérateur de la stéréotypie implicite de l'influence de l'agent sur les mesures énoncées. Ces analyses ont permis de tester nos hypothèses.

Performances au test standardisé de mathématiques. L'objectif est d'évaluer l'effet des agents virtuels pédagogiques sur les performances au test standardisé de mathématiques au temps 2, selon les trois scores de performance (score sur la globalité du test, score sur les exercices faciles, score sur les exercices difficiles). Nous avons fait l'hypothèse que l'agent avec des comportements de rapport serait associé aux meilleurs résultats en mathématiques, en particulier pour les filles. Nous n'avons pas d'hypothèse tranchée sur le genre de l'agent à cause des différents résultats trouvés dans la littérature des vrais modèles de réussite pour réduire les effets de menace du stéréotype.

L'ANOVA 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : $MER/M\overline{ER}/M\overline{\overline{ER}}$) a révélé deux effets significatifs : celui du genre de l'agent, $F(1, 306)=7.06$, $p=.008$, $\eta^2_p=.02$, ainsi que l'interaction entre le genre de l'agent et le type de l'agent, $F(2, 306)=3.47$, $p=.032$, $\eta^2_p=.02$. L'effet principal du genre de l'agent conforte l'effet obtenu au temps 1 avec la voix de la leçon : de meilleures performances étaient obtenues avec l'agent féminin ($M=50.31$, $ES=1.40$) qu'avec l'agent masculin ($M=44.86$, $ES=1.45$). L'interaction montrait que l'effet bénéfique de l'agent féminin s'observait pour l'agent avec des comportements de rapport. Comparativement à l'agent masculin avec les mêmes comportements de rapport ($M=42.13$, $ES=2.58$), l'agent féminin ($M=54.04$, $ES=2.51$) exprimant du rapport virtuel a permis aux élèves, filles comme garçons, d'obtenir de meilleurs résultats, $F(1, 306)=11.31$, $p=.001$, $\eta^2_p=.04$ (Figure 9.12).

FIGURE 9.12.

Score moyen des performances globales des élèves au temps 2, en fonction du type de l'agent et de son genre.



Note. $M\overline{ER}$: agent sans explications et sans rapport. $ME\overline{R}$: agent avec explications et sans rapport. MER : agent avec explications et avec rapport. * : différence significative. Barres d'erreur : intervalle de confiance à 95%.

L'ANOVA mixte (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : $MER/M\overline{ER}/\overline{MER}$) X 2 (difficulté des exercices : facile/difficile), avec le dernier facteur en intrasujets, n'a révélé aucun effet lié au niveau de difficulté des exercices. L'analyse a confirmé les effets des agents en intersujets sur le score global, lesquels ne dépendent donc pas du niveau de difficulté des exercices.

Sentiment de réussite en mathématiques. Comme au temps 1, l'ANOVA 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : $MER/M\overline{ER}/\overline{MER}$) a montré un effet principal du sexe de l'élève, $F(1, 305)=25.05$, $p<.001$, $\eta_p^2=.08$. Les filles ($M=3.57$, $ES=.06$) se sentaient toujours moins en réussite en mathématiques que les garçons ($M=3.97$, $ES=.06$). Un effet principal du type de l'agent, $F(2, 305)=2.93$, $p=.055$, $\eta_p^2=.02$, a montré que l'agent avec des comportements de rapport ($M=3.89$, $ES=.07$) conduisait à de meilleures auto-évaluations que l'agent sans comportements de rapport et sans explications quant à la réussite en mathématiques ($M=3.65$, $ES=.07$), $F(2, 305)=2.93$, $p=.049$, $\eta_p^2=.02$. Ce résultat est cohérent avec les résultats de nos évaluations des agents virtuels pédagogiques (Chapitre 7), où le rapport virtuel a permis aux élèves de mieux percevoir les compétences en mathématiques de l'agent, et de mieux s'identifier à lui et à sa réussite en mathématiques. En d'autres termes, le rapport virtuel a permis aux élèves de mieux percevoir les agents virtuels en tant que modèles de réussite, ce qui a pu permettre aux élèves de rapporter un plus grand sentiment de réussite en mathématiques.

Importance accordée aux mathématiques. Aucun effet significatif n'a été révélé avec l'ANOVA 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : $MER/M\overline{ER}/\overline{MER}$). L'importance accordée aux mathématiques était très élevée pour tous les élèves, indépendamment des conditions expérimentales ($M=4.34$, $ES=.04$).

9.3.2.3. Analyses de l'effet modérateur du TAI

L'objectif est d'explorer si les scores différenciés de stéréotypie implicite observés précédemment chez les filles et les garçons modéraient l'impact des agents virtuels sur leurs auto-évaluations et/ou leurs performances au temps 2. Nous avons rajouté le score du TAI en variable indépendante, en classant ces scores en deux catégories selon la médiane des scores (faible *versus* élevé). Pour chacune des analyses sur les mesures (les auto-évaluations et les performances), nous avons effectué une ANOVA 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : $MER/M\overline{ER}/\overline{MER}$) x 2 (score au TAI : faible/élevé). Les covariés étaient les mêmes que dans les analyses précédentes.

Les analyses ont révélé que le TAI a un effet modérateur de l'impact de l'agent virtuel uniquement sur l'importance accordée aux mathématiques. L'analyse de variance a montré un effet significatif de l'interaction entre toutes les variables indépendantes, $F(2, 293)=4.74$, $p=.009$, $\eta_p^2=.03$. L'analyse par contrastes avec correction de Bonferroni a montré que le niveau d'importance accordée aux mathématiques différait entre filles et garçons selon certaines conditions expérimentales. Parmi les élèves ayant un score au TAI élevé, lorsqu'ils ont interagi avec un agent féminin donnant des explications à sa réussite en mathématiques et exprimant du rapport, les filles ($M=4.47$, $ES=.19$) ont trouvé que les mathématiques étaient plus importantes comparativement aux garçons ($M=3.99$, $ES=.20$), $F(1, 293)=3.68$, $p=.056$, $\eta_p^2=.01$. Les résultats de nos évaluations des agents virtuels pédagogiques (Chapitre 7) ont non seulement montré

que le rapport virtuel permettait aux élèves de mieux percevoir les agents virtuels en tant que modèles de réussite, mais également que les élèves s'identifiaient plus à l'agent féminin qu'à l'agent masculin. Probablement qu'un tel agent virtuel a permis aux filles ayant de plus fortes associations stéréotypées de mettre un certain contrôle cognitif quant à ces associations, et de rapporter une plus grande importance aux mathématiques. Cette fois-ci parmi les élèves ayant un score plus faible au TAI, les filles (M=4.39, ES=.16) ont reporté accorder plus d'importance aux mathématiques comparativement aux garçons (M=3.71, ES=.20), après avoir interagi avec un agent féminin sans explications à la réussite en mathématiques et sans comportements de rapport, $F(1, 293)$=6.80, p=.010, η^2_p=.02. Le rapport virtuel n'est peut-être pas nécessaire pour les filles ayant peu de stéréotypie implicite (voire une contre-stéréotypie).

9.4. Discussion

L'objectif principal de cette étude était d'identifier quel serait le meilleur modèle de réussite virtuel pour réduire les effets de menace du stéréotype, afin d'augmenter les performances des filles en mathématiques, sans pour autant baisser celles des garçons. Nous avions fait l'hypothèse que ce serait l'agent avec des comportements de rapport virtuel, et expliquant sa réussite en mathématiques par les efforts et le travail, qui réduirait le plus les effets de menace du stéréotype. Notre hypothèse sur le genre n'était pas tranchée, dans la mesure où plusieurs études ont montré un effet bénéfique uniquement avec un modèle de réussite incarné par une femme (Leroy et al., 2022 ; Marx & Roman, 2002 ; McIntyre et al., 2011), mais que d'autres études ont montré que c'était finalement l'explication de la réussite par le travail et les efforts qui primait (Bagès & Martinot, 2011 ; Bagès et al., 2016). Cependant, ces dernières études ne proposaient pas de composante visuelle, car les élèves devaient uniquement lire une histoire personnelle sur des modèles de réussite. Peut-être que le genre du modèle redevient important lorsqu'il est possible d'interagir avec un modèle expliquant sa réussite par son travail.

Pour tester ces hypothèses, nous avons créé plusieurs agents virtuels pédagogiques dont le rôle principal était d'expliquer les notions clés d'une leçon de mathématiques (le parallélisme chez les 6e, et les priorités opératoires chez les 5e). Dans cette étude, nous avons procédé en deux phases. La première phase permettait de recueillir les auto-évaluations des élèves en mathématiques, une première mesure de leur degré de stéréotypie implicite, et de les habituer à la passation d'un test de mathématiques en numérique. Cette passation numérique se déroulait sans la présence d'un agent virtuel, mais avec une voix pré-enregistrée et générée automatiquement (par Text-to-Speech) qui expliquait les notions clés de la leçon en mathématiques. La deuxième phase permettait de tester l'efficacité de nos agents virtuels. Elle se déroulait en présence d'un agent virtuel pédagogique, qui se présentait et expliquait la leçon de mathématiques, suivant le plan expérimental intersujet 2 (sexe de l'élève : fille/garçon) X 2 (genre de l'agent : féminin/masculin) X 3 (type de l'agent : MER/$ME\overline{R}$/$M\overline{ER}$).

L'agent féminin exprimant des comportements de rapport virtuel a été bénéfique pour les filles comme pour les garçons, ce qui invalide notre hypothèse : il n'a pas été plus bénéfique pour les filles. Est-ce à dire que notre agent n'a pas réduit les effets de menace du stéréotype ? Il est probablement compliqué de parler de réduction de menace du stéréotype, dans la mesure où il n'y avait déjà aucune différence de performance au temps 1, pouvant être considéré comme une *baseline*. Les tests semblaient pourtant suffisamment difficiles : les élèves ont

obtenu un score de 42.92% au temps 1, et 47.59% au temps 2. Des résultats intéressants ont tout de même été trouvés : la voix féminine de la vidéo de la leçon a permis aux élèves d'obtenir de meilleurs résultats, comparativement à la voix masculine. En particulier, la voix féminine permettait aux filles d'obtenir de meilleurs résultats sur les questions difficiles. Même s'il n'y a pas eu de différence de performance entre filles et garçons dans la condition de la voix masculine, on peut se demander si cette simple voix féminine a réduit les effets de menace du stéréotype, car c'est sur les questions difficiles que les effets de menace du stéréotype s'observent le plus. Dans tous les cas, l'agent féminin avec des comportements de rapport et des explications de la réussite par le travail a eu un effet bénéfique pour tous les élèves, quel que soit leur genre.

Le rapport virtuel a non seulement des effets bénéfiques sur les performances en mathématiques (du moins lorsqu'il est exprimé par un agent féminin), mais également sur les auto-évaluations en mathématiques. L'agent avec des comportements de rapport virtuel, cette fois-ci indépendamment de son genre, a conduit à un plus grand sentiment de réussite en mathématiques. Pour mieux comprendre ce résultat, certains résultats obtenus lors de nos pré-tests des agents virtuels peuvent nous aider. Le rapport virtuel permettait de faire de nos agents de meilleurs modèles de réussite (Chapitre 7). Les élèves se sont plus identifiés aux agents exprimant du rapport virtuel et à leur réussite en mathématiques, et les ont trouvés plus compétents en mathématiques. Il est possible que le rapport virtuel de nos agents, couplé aux explications de la réussite en mathématiques par le travail, a incité les élèves à s'identifier à leur réussite, et donc à reporter un plus grand sentiment de réussite en mathématiques et de meilleures performances.

Concernant la stéréotypie implicite, celle-ci a eu un effet modérateur de l'influence des agents, uniquement sur le niveau d'importance accordée aux mathématiques. Ainsi, pour les élèves ayant un haut score au TAI, l'agent féminin avec des comportements de rapport virtuel et avec des explications de la réussite par le travail a permis aux filles de trouver les mathématiques plus importantes comparativement aux garçons. Ce résultat se rapproche de celui de FRANCESCHINI et al. (2014), où une consigne mettant en doute le stéréotype (e.g., "les femmes ont mieux réussi que les hommes dans ce test") permettait uniquement aux femmes avec une stéréotypie implicite élevée de mieux réussir un test de mathématiques. Pour les élèves avec une plus faible stéréotypie implicite, les filles ont reporté accorder plus d'importance en mathématiques par rapport aux garçons, seulement après avoir interagi avec l'agent virtuel féminin ne donnant aucune explication quant à sa réussite, et avec aucun comportement de rapport. Comme pour l'étude de KIEFER et SEKAQUAPTEWA (2007), il semblerait qu'il ne soit pas nécessaire de remettre le stéréotype en doute pour les filles qui n'associent pas (ou peu) les garçons avec les mathématiques.

En résumé, la propension de l'agent à établir un rapport virtuel dans les interactions avec les élèves semble être une caractéristique fondamentale pour obtenir un effet bénéfique des agents pédagogiques, pour les filles comme pour les garçons. Le genre de l'agent exprimant du rapport virtuel n'a pas eu d'importance sur les auto-évaluations. Cependant, l'effet bénéfique du rapport virtuel s'exprime notamment par un agent féminin sur les performances en mathématiques. Ceci est cohérent avec les résultats du temps 1, où la simple voix féminine de la vidéo permettait aux filles de mieux réussir les exercices difficiles. Finalement, est-ce qu'une voix féminine ne serait pas suffisante pour aider les jeunes filles en mathématiques ? Il est

difficile de trancher sur cette question-là. Les évaluations perceptives de nos agents (Chapitre 7) semblent indiquer que le rapport virtuel permet d'augmenter l'identification à l'agent, en le percevant comme plus compétent en mathématiques. On pourrait assimiler la voix à l'agent sans comportements de rapport virtuel. Cependant, nous n'avons tout de même pas mesuré le degré d'identification à la voix, ni le sentiment de rapport virtuel ressenti avec la voix, d'où le fait que nous ne pouvons pas complètement trancher.

Des résultats inattendus, hors des objectifs initiaux mais très intéressants, ont été trouvés concernant les différentes notations utilisées par les enseignants dans leur classe (classique *versus* couleur). Cette différence de notation semble avoir un impact à différents niveaux. Premièrement, le sentiment de réussite des filles en mathématiques était moins élevé que celui des garçons dans les classes où les enseignants utilisent la notation classique. La notation par couleur éliminerait les différences filles/garçons sur l'auto-évaluation en mathématiques, où les filles se sentaient autant en réussite que les garçons. Deuxièmement, comparativement aux garçons qui présentaient une stéréotypie implicite de genre en mathématiques, le score obtenu par les filles traduisait une absence de stéréotypie implicite, voire une contre-stéréotypie dans les classes où les enseignants utilisent la notation par couleur. Cette notation réduirait également le niveau de stéréotypie implicite chez les garçons. Les comparaisons sociales probablement encouragées par le système de notation classique étant au cœur des disparités entre les filles et les garçons et la formation des représentations stéréotypiques, la notation par couleur permettrait d'échapper à ces effets. Ces différences de niveau de stéréotypie implicite selon la notation utilisée montre qu'il reste malléable chez les enfants/pré-adolescents de cet âge-là, filles comme garçons. Ceci est consistent avec l'étude de GALDI et al. (2014) montrant que la stéréotypie implicite était malléable pour les filles de 6 ans. Manifestement, cette malléabilité est toujours présente en 6e et en 5e. Ceci encourage à développer au niveau du collège les actions destinées à réduire les effets de menace du stéréotype.

Cependant, l'utilisation de la notation par couleur était associée à des scores plus bas dans nos tests de mathématiques standardisés, comparativement à l'évaluation classique basée sur des notes. Cette tendance est probablement due au fait que la notation par couleur pourrait être moins propice à la comparaison entre élèves, ce qui est pourtant essentiel pour évaluer leur niveau de compétence et stimuler leur motivation à s'améliorer. Une tension éthique entre deux objectifs se pose quant à l'utilisation des différentes notations dans les classes : une meilleure performance en mathématiques, où nos résultats tendent à favoriser la notation classique, et la réduction de stéréotypie implicite ainsi qu'un pas vers l'égalité entre filles et garçons sur leurs auto-évaluations en mathématiques, où nos résultats recommandent plutôt la notation par couleur. Il faut cependant noter que ces résultats ont été observés sur notre échantillon, qui est composé principalement d'élèves scolarisés en REP+. Il serait intéressant de tester ces effets potentiellement différenciés de la notation classique et par couleur plus largement, en incluant des élèves non-REP. Ceci serait possible en analysant les performances en mathématiques aux tests standardisés tels que le brevet. De plus, nous ne connaissons pas les conséquences de la notation en couleur au collège ni sur les performances, ni sur les auto-évaluations, ni sur la stéréotypie implicite dans le long terme, lorsque ces élèves devront faire face au système de notation classique dans la suite de leurs études.

Il est important de souligner que certaines limites dans cette étude sont à discuter. Premièrement, il y a eu un biais de répartition dans les conditions expérimentales, où des élèves

meilleurs que d'autres étaient présents dans certaines conditions. Nous l'avons pris en compte en covarié. Cette limite a finalement été amoindrie, dans la mesure où les élèves avec les meilleures moyennes trimestrielles en mathématiques étaient dans la condition de la voix masculine et de l'agent masculin, et que les élèves ayant obtenu les meilleures notes à nos tests standardisés étaient dans la condition inverse. Deuxièmement, la notation par couleur utilisée dans certaines classes a pu empêcher de faire ressortir les effets de menace du stéréotype, en effaçant les différences de performance entre les filles et les garçons. Cela ne veut pas forcément dire que la menace du stéréotype n'agit pas, mais qu'il est plus difficile de l'observer. De plus, la conversion de ces couleurs en note plus classique fournie par les enseignants concernés est loin d'être optimale. Cela contribue inévitablement à écraser des différences inter-individuelles entre les élèves, et donc entre les établissements concernés. Une autre limite concernant les profils des participants a également pu masquer les effets de menace du stéréotype. En effet, une majorité de classes en REP+ a participé à l'étude. Cette volonté s'est fondée sur la nécessité de tester les agents virtuels pédagogiques auprès d'élèves moins favorisés, qui auraient le plus besoin de dispositif dans les classes. Néanmoins, cela comportait un risque car il est plus aisé d'observer les effets de menace du stéréotype chez une population plus favorisée. Cela s'explique par le désengagement scolaire plus fort chez les garçons que chez les filles dans les milieux en REP+ (BREDA et al., 2024). L'effet de menace du stéréotype peut donc être masqué par la chute des performances des garçons dû à ce désengagement scolaire. Les dernières limites concernent les passations de l'étude sur le terrain. Les classes d'un même établissement et d'un même niveau (e.g., 6e ou 5e) ont effectué les passations en parallèle pour éviter que les élèves ayant déjà passé l'étude puissent en parler plus en détail à leurs camarades d'autres classes. L'expérimentatrice a donc pris part à certaines passations des évaluations dans leur intégralité, laissant quelques classes être surveillées uniquement par un enseignant (pas forcément de mathématiques), parfois par un assistant d'éducation. Par conséquent, certains biais ont pu être présents (ajout de consignes non-sollicitées, libertés avec le dispositif, etc). L'expérimentatrice était présente dans chaque établissement lors des passations, et a pu discuter avec le personnel en charge de la surveillance des classes pour avoir un retour de terrain. Néanmoins, nous ne pouvons être sûres que l'ensemble de ces classes, en incluant principalement la personne en charge de la surveillance, a participé sérieusement à l'étude. De plus, l'expérimentatrice s'étant présentée dans chaque classe pour expliquer les consignes (y compris les classes où elle n'était pas en charge de la surveillance complète), cela inclus inévitablement un biais. Des études ont montré que le genre de l'expérimentateur pouvait avoir un effet sur la menace du stéréotype, où les filles pouvaient percevoir l'expérimentatrice comme un modèle de réussite, ce qui améliorait leurs résultats comparativement aux filles dont les consignes ont été expliquées par un expérimentateur. Enfin, en retirant les absents et les tricheurs, nous avons un plus petit nombre de données exploitables au temps 2 (Tableau 9.6), alors qu'il y avait plus de 500 élèves inscrits pour participer (Section 9.2.1).

Malgré ces limites, cette étude apporte des contributions à la fois à un niveau pratique et théorique à la littérature scientifique. Tout d'abord, nos résultats permettent d'approfondir les profils des modèles de réussite efficaces dans les apprentissages. Des études ont montré qu'expliquer la réussite par le travail et les efforts permettaient de rendre un modèle accessible, ce qui permet de s'y identifier (MCINTYRE et al., 2011) et de réduire les effets de menace du stéréotype (BAGÈS et al., 2016). L'explication de l'intelligence malléable, sans forcément être donnée par un modèle de réussite, est bénéfique pour tous les élèves, filles comme garçons

(BLACKWELL et al., 2007). Notre travail vient enrichir ces résultats, en montrant qu'un agent virtuel ayant des comportements de rapport virtuel couplé à ces explications permet d'augmenter le sentiment de réussite en mathématiques - bien que l'on observe toujours un écart entre les filles et les garçons, où les filles se sentent tout de même moins en réussite que les garçons malgré leur performance similaire, un résultat bien connu dans la littérature en sciences sociales (CATSAMBIS, 1994 ; ZANDER et al., 2020). Plus précisément, lorsque ce rapport virtuel a été exprimé par un agent féminin, cela a permis d'augmenter les performances en mathématiques, et ce pour les filles comme pour les garçons. Le rapport virtuel est un élément important dans l'apprentissage, permettant de percevoir plus facilement un agent comme un modèle de réussite (résultats de nos études pilotes au Chapitre 7). Le rapport permet également de rendre une classe plus harmonieuse. Les enseignants ayant des comportements de rapport permettent en effet de créer un climat de classe confortable où les élèves peuvent mieux apprendre (TSUI, 1996), où les étudiants sont plus susceptibles d'assister aux cours, d'être attentifs, d'apprécier le sujet du cours BENSON et al. (2005), et où les étudiants participent plus en classe et obtiennent de meilleurs résultats scolaires (FRISBY & MARTIN, 2010 ; GRANITZ et al., 2009). Nous avons répliqué l'amélioration des résultats scolaires, ici en mathématiques avec des enfants et avec le rapport virtuel. On peut cependant se demander si cette plus grande performance et ce plus grand sentiment de réussite serait dû au rapport virtuel couplé avec un certain sentiment d'anonymat que les individus peuvent éprouver lors d'une interaction avec un agent virtuel, à priori non contrôlé par un humain, faisant preuve de comportements de rapport. Le rapport virtuel permet aux individus de plus se confier (LUCAS et al., 2017), et le sentiment d'anonymat émanant de l'interaction avec un agent virtuel ayant des comportements de rapport virtuel permet de ressentir moins de jugements (DEVAULT et al., 2014). Peut-être que cela a enlevé une certaine pression que peuvent ressentir les élèves, augmentant ainsi leur performance et leur sentiment de réussite. De plus, bien que notre agent virtuel avec rapport s'adapte aux réponses de l'élève lorsque l'agent lui pose une question, celui-ci reste scripté pour que tous les élèves dans cette condition aient une expérience similaire. Cela peut être vu comme une limite, car les élèves n'ont pas pu discuter réellement avec l'agent. Néanmoins, des études ont montré que les apprenants supposent que des agents scriptés fournissent des réponses dynamiques et adaptatives, même si ce n'est pas le cas (BAYLOR & CHANG, 2002). Il est intéressant de noter que même un agent scripté avec des comportements de rapport et quelques adaptations à l'élève, sans pouvoir prendre en compte les signaux des utilisateurs, a un impact positif sur les performances des élèves. En résumé, le rapport virtuel apporte une plus-value aux explications de la réussite du modèle de réussite par le travail.

Nous apportons d'autres contributions au niveau théorique, cette fois-ci sur le genre de nos agents virtuels pédagogiques. Quelques-uns de nos résultats sont en cohérence avec ceux de BAGÈS et MARTINOT (2011). Notamment, les performances des garçons comme pour les filles étaient meilleures avec un modèle de réussite féminin. Cependant, dans les études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016), lorsque le modèle expliquait qu'il a eu de bons résultats en mathématiques grâce à son travail et ses efforts, les performances des apprenants étaient encore meilleures, indépendamment du genre du modèle et de l'apprenant. Contrairement aux études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016) où le genre d'un modèle de réussite n'avait pas d'importance dès lors qu'il est expliqué que ce modèle a de bons résultats en mathématiques grâce à son travail, c'est l'agent virtuel féminin avec des comportements de rapport qui a permis d'obtenir les meilleurs résultats pour les filles et les

garçons, malgré la présence d'explication sur l'intelligence malléable. Nos résultats suggèrent que la féminité joue un rôle important dans l'augmentation des résultats des élèves, et plus particulièrement des filles. En effet, dans la première phase de notre étude lorsqu'une vidéo avec une voix off rappelait les notions clés de la leçon de mathématiques, c'est la voix féminine comparativement à la voix masculine qui a permis à tous les élèves d'obtenir de meilleurs résultats, et qui a permis aux filles de produire de meilleure performance dans les exercices difficiles. Contrairement aux résultats de MARX et ROMAN (2002), où les performances des hommes ont baissé avec un modèle de réussite féminin très compétent en mathématiques, les garçons de notre étude ont mieux réussi les tests de mathématiques avec la voix féminine au temps 1, et avec un agent féminin au temps 2 (tout comme les filles). Les hommes participants de l'étude de MARX et ROMAN (2002) ont cependant mieux réussi avec un modèle de réussite féminin modérément compétent en mathématiques, car celui-ci défie moins les stéréotypes de genre. Or, il est peu probable que l'agent féminin avec des comportements de rapport a été perçu comme modérément compétent en mathématiques, puisque nos études pilotes 1A et 1B (Chapitre 7) montrent toutes les deux que tous les élèves perçoivent les agents féminins et masculins comme étant très compétents en mathématiques, et que le rapport virtuel augmentait la perception de la compétence en mathématiques. Nous avions tout d'abord pensé que les garçons auraient pu percevoir les agents féminins comme moins compétents qu'eux en mathématiques, ce qui aurait donné un sentiment de facilité ("si elle peut le faire, alors moi aussi"), mais comme nous venons de le rappeler, les agents féminins ont été perçus comme très compétents en mathématiques (un score plus élevé que le sentiment de réussite en mathématiques des garçons). Une interprétation plus probable est de lier la féminité avec le sentiment de rapport et le niveau d'anthropomorphisme. En effet, une récente étude a montré que les femmes et les robots féminins sont perçus comme plus humains que leur homologues masculins (BORAU et al., 2021). Dans notre étude pilote s'approchant le plus de la configuration de cette étude (étude pilote 1B, Chapitre 7), l'agent féminin a été significativement perçu comme générant plus de rapport virtuel que l'agent masculin. De plus, dans cette même étude pilote, l'agent féminin avec des comportements de rapport virtuel a été jugé plus proche de l'humain que ce même agent féminin sans comportements de rapport virtuel. Dans l'étude principale, peut-être qu'un sentiment de rapport plus fort a été perçu avec l'agent féminin avec des comportements de rapport virtuel, résultant ainsi à une plus grande perception d'anthropomorphisme. Ce serait ces trois facteurs qui expliqueraient pourquoi l'agent féminin avec des comportements de rapport a permis à tous les élèves d'obtenir les meilleurs résultats au test standardisé de mathématiques. Cela expliquerait également pourquoi, contrairement aux études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016), le genre du modèle de réussite a une importance. Cette interprétation est cependant liée à une problématique d'éthique que nous allons davantage développer dans la partie "Discussion Générale".

Enfin, cette étude apporte également des contributions au niveau pratique, notamment par la création d'un environnement virtuel d'apprentissage *safe* pour l'identité des filles. L'exposition répétée à un modèle de réussite féminin (l'agent virtuel pédagogique féminin avec comportements de rapport virtuel) pourrait permettre aux filles de se sentir de plus en plus en confiance, et aux garçons de voir des filles en sciences, ce qui pourrait réduire les stéréotypes de genre. Cela pourrait également réconcilier les résultats que nous avons trouvés sur les différences qu'engendrent le type de notation des enseignants. La notation par couleur (e.g., associer l'élève à un groupe de compétence plutôt qu'à une note) éliminerait les différences filles/garçons en mathématiques, et les élèves présentaient moins de stéréotypies implicites

(en particulier chez les filles où le score obtenu traduisait une absence de stéréotypie implicite, voire une contre-stéréotypie). Cependant, l'utilisation de la notation en couleur est liée à des scores plus bas en mathématiques, comparativement aux élèves notés de manière plus classique, avec une note allant de 0 à 20. Les scores de stéréotypie implicite n'étant pas très élevés chez l'ensemble de nos participants, et nos résultats montrant que la stéréotypie implicite reste malléable à cet âge-là, l'exposition continue de nos participants à notre agent virtuel féminin avec des comportements de rapport pourrait alors créer de nouveaux automatismes de stéréotypies implicites, où les élèves associeraient les garçons et les filles aux mathématiques, tout en améliorant leurs performances dans cette matière. Plus de perspectives sont énoncées dans la partie "Discussion Générale".

9.5. Points clés du chapitre

Résultats clés.

Temps 1 (explication de la leçon par une voix).

- La voix féminine de la vidéo de la leçon était associée aux meilleurs résultats en mathématiques comparativement à la voix masculine.
- Cette voix féminine a permis aux filles de mieux réussir les exercices difficiles, et aux garçons de mieux réussir les exercices plus faciles.

Temps 2 (explication de la leçon par un agent virtuel pédagogique).

- L'agent féminin avec des comportements de rapport virtuel a permis à tous les élèves de mieux réussir le test de mathématiques.
- L'agent avec des comportements de rapport virtuel, indépendamment de son genre, a conduit à de plus grand sentiment de réussite en mathématiques.
- Le rapport virtuel est nécessaire pour que les filles qui ont un degré de stéréotypie implicite élevé accordent plus d'importance aux mathématiques.

Temps 1 et 2.

- Pas de différence filles/garçons dans les performances en mathématiques et dans l'importance accordée aux mathématiques.
- Les filles se sentaient moins en réussite en mathématiques comparativement aux garçons.

Résultats secondaires.

- La notation par couleur permettrait aux filles de se sentir autant en réussite en mathématiques que les garçons, contrairement à la notation classique (e.g., avec une note) où elles se sentent moins en réussite.
- La notation par couleur diminuerait la stéréotypie implicite, et plus particulièrement chez les filles.
- La notation par couleur diminuerait les performances en mathématiques de tous les élèves.

Troisième partie

Discussion générale, limites, perspectives

10. Discussion générale

Sommaire

10.1. Principales contributions de la thèse 257
10.2. Implications théoriques 260
10.3. Implications pratiques 262
10.4. Implications éthiques 263
10.4.1. Apparence des agents virtuels pédagogiques 264
10.4.2. Comportements verbaux et non-verbaux des agents virtuels pédagogiques 265
10.4.3. Quelques réflexions personnelles 267
10.5. Limites 268
10.5.1. Limites des expérimentations 268
10.5.2. Limites des agents virtuels pédagogiques 269
10.6. Perspectives 270

10.1. Principales contributions de la thèse

L'objectif principal de cette thèse était de créer des modèles de réussite à travers des agents virtuels pédagogiques auxquels les enfants peuvent s'identifier afin de réduire les effets de menace de stéréotype, et donc améliorer les performances des jeunes filles de 6e et de 5e. Un tel agent virtuel pédagogique pourrait servir d'outil pédagogique complémentaire aux enseignants, qui pourraient l'utiliser par exemple pour occuper une partie des élèves pendant qu'ils se concentrent sur d'autres élèves qui rencontrent des difficultés. Cet agent pourrait également être utilisé comme support pour les devoirs à la maison.

La principale contribution de cette thèse dans le domaine des agents virtuels pédagogiques est la création d'agents virtuels qui simulent des modèles de réussite auxquels les élèves peuvent s'identifier. Pour cela, nous avons d'une part extrait les caractéristiques des modèles de réussite à partir de la littérature en psychologie sociale sur les profils de modèles efficaces contre la menace du stéréotype (Section 2.3.4.3). D'autre part, notre modèle computationnel de comportements verbaux et non-verbaux des agents virtuels repose sur les théories dites du rapport (Section 3.4), et est décrit à travers des scripts à partir de balises de style XML pour simuler une interaction de rapport virtuel avec l'élève. Ce modèle computationnel ainsi que les agents ont été validés dans nos études pilotes, conduites hors-laboratoire, avec des enfants dans les classes. Nos résultats ont notamment montré que le modèle computationnel de rapport virtuel implémenté dans nos agents permettait aux élèves de mieux les percevoir comme des modèles de réussite. En particulier, les élèves ont estimé que ces agents étaient plus compétents en mathématiques, et se sont plus identifiés à eux et à leur réussite en mathématiques, comparativement aux mêmes agents dépourvus du modèle computationnel de rapport (e.g., moins de mouvements, pas d'expressions faciales positives). **Ces résultats valident l'efficacité des agents que nous avons développés et du modèle computationnel sous-jacent pour simuler des modèles de réussite auxquels les élèves peuvent s'identifier**. Cette contribution constitue une avancée significative dans le domaine des agents virtuels, car il n'existait pas à notre connaissance de modèle computationnel de modèle de réussite pour les agents virtuels.

Les travaux de recherche menés dans cette thèse offrent également un nouvel éclairage sur la perception stéréotypée des agents par les enfants, ce qui constitue une deuxième contribution dans le domaine des agents virtuels. Nos résultats suggèrent que les enfants n'auraient pas les mêmes perceptions des stéréotypes que les adultes. À noter cependant que nous n'avons pas inclus de participants adultes dans nos études. Nous comparons simplement nos résultats obtenus avec des enfants à ceux obtenus dans d'autres études conduites avec des adultes et donc avec des agents virtuels différents des nôtres. Nos études montrent en effet que les élèves n'ont pas été influencés par les stéréotypes de genre dans leur perception du niveau de compétence en mathématiques des agents, contrairement certaines études réalisées avec des participants adultes, où les agents féminins sont généralement perçus comme moins compétents que les agents masculins (BAYLOR & KIM, 2004 ; NUNAMAKER et al., 2011). On observe tout de même quelques éléments stéréotypiques dans la perception des élèves. Notamment, le niveau d'anthropomorphisme des agents dépendait de leur genre et de leurs comportements : les élèves ont eu tendance à percevoir un personnage virtuel féminin comme plus réaliste lorsqu'il adoptait un comportement socio-émotionnel positif avec des sourires, que lorsque ces signaux étaient absents. Il est intéressant de noter que c'était l'inverse pour l'agent masculin. Ces résultats sont cohérents avec les recherches qui montrent que, conformément aux attentes

sociétales envers les femmes, les individus s'attendent à ce que les femmes sourient plus que les hommes (DEUTSCH et al., 1987 ; LAFRANCE & HECHT, 1995). Par conséquent, un écart par rapport à ce comportement attendu influence négativement la perception des femmes qui ne sourient pas, et l'absence de sourire peut être préjudiciable à l'image d'une femme par rapport à un homme. **En résumé, ces travaux de thèse ont permis de développer des agents virtuels pédagogiques auxquels les enfants peuvent s'identifier. Ces agents virtuels ont été validés dans des conditions réelles d'utilisation en classe. De plus, nos études perceptives ont mis en lumière les stéréotypes que les enfants peuvent attribuer aux agents virtuels pédagogiques. Ces résultats ouvrent de nombreuses perspectives pour améliorer l'apprentissage des enfants à travers ce type d'agents virtuels pédagogiques (Section 10.6).**

En ce qui concerne notre contribution à la littérature sur la menace du stéréotype, l'absence de différence de performance entre les filles et les garçons sur nos tests standardisés de mathématiques soulève une question importante : nos agents ont-ils permis de réduire les effets de menace du stéréotype chez les filles en mathématiques ? Plusieurs éléments semblent indiquer que non. Tout d'abord, nous n'avons observé aucune différence de performance entre les filles et les garçons aux tests de mathématiques, que les tests aient été réalisés avec ou sans l'intervention d'un agent (comparaison entre le temps 1 et le temps 2). Ensuite, nos agents n'ont pas influencé différemment les performances des filles et des garçons. L'agent virtuel pédagogique féminin, incarnant un modèle de réussite et doté de comportements de rapport, a permis en effet d'améliorer les performances de tous les élèves, indépendamment de leur genre. Enfin, aucun agent n'a permis de booster le sentiment de réussite en mathématiques des filles (et donc de réduire l'écart observé à la faveur des garçons), contrairement à ce que nous aurions pu attendre dans le cas d'une réduction de la menace de stéréotype.

Au-delà de la question de l'efficacité de nos agents pour réduire l'effet de menace du stéréotype, se pose la question de la présence de cet effet au sein de notre échantillon. Compte tenu du niveau de difficulté des tests standardisés proposés aux élèves, nous aurions dû observer une différence de performance à la faveur des garçons au temps 1, lorsque la passation des tests s'est déroulée sans intervention des agents virtuels (c'est-à-dire en situation supposée de menace du stéréotype). Or, cela n'a pas été le cas. Est-ce à dire qu'il n'y a pas d'effet de menace du stéréotype dans notre échantillon ? Pas forcément. L'absence de différence de performance entre les filles et les garçons peut s'expliquer par certaines caractéristiques de nos participants, en lien avec le type de collèges dans lesquels ils sont scolarisés. La grande majorité de nos participants est en effet issue de collèges REP+ (249 sur 326 données d'élèves analysées). Ces collèges se situent dans des quartiers ou secteurs isolés connaissant les plus grandes concentrations de difficultés sociales ayant des incidences fortes sur la réussite scolaire. Or, il est plus difficile d'observer des effets de menace du stéréotype dans ces milieux, en raison du désengagement scolaire plus fort chez les garçons que chez les filles (BREDA et al., 2024). L'effet de menace du stéréotype a pu être masqué par les moins bonnes performances des garçons en raison de leur désengagement scolaire. Cela ne veut pas nécessairement dire qu'il n'existe pas. Par ailleurs, il se pourrait également que certaines modalités des conditions de passation au temps 1 aient contribué, de façon inattendue, à cette égalité de performances entre filles et garçons, en réduisant les effets de menace du stéréotype chez les filles. En effet, la leçon et le test de mathématiques étaient accompagnés d'une voix off, qui était soit féminine soit masculine. L'utilisation de la voix avait uniquement pour objectif de standardiser les conditions de révision de la leçon et pallier les différences éventuelles de capacité de lecture des élèves. Or, le genre de la voix a eu une influence : sur les exercices les plus difficiles, les filles ayant

entendu la voix féminine ont obtenu les meilleurs résultats, alors que celles ayant entendu la voix masculine ont obtenu les moins bons résultats. Autrement dit, les filles ont bénéficié de la voix féminine sur les exercices difficiles, c'est-à-dire là où l'on observe habituellement les effets de menace du stéréotype. Pris ensembles, ces éléments nous conduisent à nuancer nos conclusions précédentes. Le fait que l'agent féminin avec des comportements de rapport virtuel (et avec les explications de la réussite par le travail) ait bénéficié aux filles comme aux garçons pourrait refléter : (1) une réduction de la menace du stéréotype chez les filles, et (2) une aide nécessaire aux garçons en décrochage scolaire.

Enfin, cette thèse apporte des contributions en psychologie sociale s'agissant de l'utilisation du Test des Associations Implicites (TAI) auprès des enfants/pré-adolescents. Nous avons traduit en français et recodé le TAI genre-mathématiques-lecture de CVENCEK et al. (2011) en HTML, JavaScript, et PHP pour la version web (avec les mots lus par une voix lorsqu'ils apparaissent à l'écran). Le code est disponible en open-source sur https://github.com/ArmandoMarjorie/French_child_IAT pour que les chercheurs puissent l'utiliser. Notre version est facilement modifiable pour utiliser un TAI autre que genre-mathématiques-lecture en changeant les mots, les audios des mots, et les catégories. Il est à noter qu'aucune version française du TAI pour enfants/pré-adolescents n'était disponible jusqu'à présent. Nous avons élaboré les consignes et instructions en collaboration avec l'équipe d'enseignants de mathématiques d'un collège de REP+ et quelques enfants de 6e de notre entourage pour qu'elles soient le aussi compréhensibles que possible. Cependant, nos résultats ainsi que nos observations dans les classes ont révélé que les élèves n'ont pas pleinement compris la tâche lors du premier essai (temps 1), et l'ont mieux comprise lorsqu'ils l'ont repassé au deuxième temps. Lorsqu'ils ont découvert la tâche la première fois, ils se sont davantage focalisés sur la performance que sur la vitesse, contrairement à l'objectif du TAI qui est de répondre le plus rapidement possible en faisant le moins d'erreur possible. Cette divergence pourrait s'expliquer en partie par les consignes habituellement données par les enseignants en classe, qui encouragent les élèves à prioriser l'exactitude sur la vitesse de réponse (e.g., prendre son temps avant de répondre, se relire plusieurs fois). Par conséquent, seuls les scores obtenus au deuxième temps de mesure du TAI ont été utilisés pour nos analyses, et en prenant soin de ne retenir que les élèves qui avaient bien passé le test deux fois.

Deux résultats impliquant le TAI méritent attention. Premièrement, la stéréotypie implicite semble modérer l'effet du type d'agents virtuels sur l'importance accordée par les élèves aux mathématiques. Ainsi, après avoir interagi avec l'agent féminin exprimant des comportements de rapport, les filles accordent plus d'importance que les garçons aux mathématiques uniquement si elles ont un niveau élevé de stéréotypie implicite. Quant aux filles ayant une stéréotypie implicite moins élevée, c'est l'agent féminin n'exprimant pas de rapport et n'expliquant pas sa réussite en mathématiques qui leur a permis de trouver les mathématiques plus importantes, comparativement aux garçons. Ceci est en cohérence avec les rares études sur l'effet modérateur de la stéréotypie implicite sur l'effet de menace du stéréotype (FRANCESCHINI et al., 2014 ; KIEFER & SEKAQUAPTEWA, 2007). Deuxièmement, l'utilisation d'une notation par couleur (au lieu d'une note sur 20) dans les classes s'est avérée associée à un moindre niveau de stéréotypie implicite chez les élèves, en particulier chez les filles, suggérant l'idée que la stéréotypie implicite serait encore malléable à cet âge-là. Il est important de souligner que les quelques études qui ont utilisé le TAI chez les enfants ou les pré-adolescents ne précisent pas comment ce test est présenté aux enfants, ne mentionnent pas leur niveau de compréhension des consignes, n'ont pas réalisé de test-retest, et ne fournissent pas les temps de réponses

moyens pour chaque étape du TAI. Il ne nous a donc pas été possible de comparer nos résultats à ceux de ces études antérieures.

En résumé, nous avons proposé une version française du TAI genre-mathématiques-lecture adaptée aux enfants et pré-adolescents, pouvant être facilement utilisée et/ou modifiée pour créer d'autres TAI. Nos travaux apportent des informations nouvelles sur l'impact que peut avoir la stéréotypie implicite sur l'influence des agents virtuels concernant les auto-évaluations des élèves en mathématiques. Nos travaux montrent de plus que la stéréotypie implicite reste encore malléable à cet âge-là.

10.2. Implications théoriques

Cette thèse présente plusieurs implications théoriques pour la littérature sur la menace du stéréotype en psychologie sociale, et pour celle sur les agents virtuels en informatique. La première partie de la thèse montre tout d'abord que l'on peut nourrir ces deux disciplines l'une de l'autre pour créer des agents virtuels pédagogiques afin d'améliorer l'apprentissage des élèves. Nous avons combiné les études de ces littératures dans la deuxième partie de cette thèse : la littérature sur la menace du stéréotype permet d'identifier précisément les caractéristiques visuelles et comportementales nécessaires pour créer des agents susceptibles de réduire les effets des stéréotypes de genre, et la littérature sur les agents virtuels pédagogiques apporte la notion du modèle computationnel de rapport virtuel dans les interactions entre un modèle de réussite en mathématiques et un apprenant, pouvant être transposé dans les interactions interhumaines. Dans ce contexte interdisciplinaire, nous avons ainsi créé des agents virtuels pédagogiques, une fille et un garçon, incarnant un modèle de réussite en mathématiques et exprimant des comportements de rapport virtuel. Cette construction des agents, fondée à la fois sur les études sur la menace du stéréotype et sur les agents virtuels pédagogiques, a permis d'éviter des rôles stéréotypés (e.g., filles peu compétentes en STIMs, assistantes virtuelles). Il est à noter que la littérature sur les agents virtuels et les assistants vocaux montre souvent une sélection du genre de ces entités virtuelles sur la base de leur rôle stéréotypé (West et al., 2019), ce que nous avons cherché à éviter dans notre approche.

Nos résultats soulignent le rôle crucial du rapport virtuel dans le processus l'apprentissage. Tout d'abord, les élèves se sont davantage identifiés aux agents dotés de notre modèle computationnel de rapport, les percevant comme plus compétents en mathématiques. Cela apporte un nouvel éclairage sur la façon dont un modèle de réussite incarné par un humain pourrait se comporter. Des études antérieures ont montré qu'en situation réelle, les élèves ne s'identifiaient pas aux femmes présentées comme des modèles de réussite lorsque celles-ci exprimaient de l'anxiété quant à leurs capacités en mathématiques (Beilock et al., 2010). Nos résultats suggèrent qu'une plus grande présence d'enseignantes passionnées par les mathématiques, qui manifestent leur intérêt et démontrent leur compétence dans ce domaine, pourrait susciter une identification plus forte de la part des élèves. Nous faisons le parallèle avec le monde réel car les individus ont une propension à interagir avec des agents virtuels comme s'ils étaient humains (Heyselaar et al., 2023 ; Nass & Moon, 2000). Nos résultats permettent donc d'approfondir les profils des modèles de réussite. D'autres recherches ont mis en évidence l'efficacité de l'explication de la réussite par le travail et les efforts, ce qui rend un modèle plus accessible et favorise l'identification avec celui-ci (Bagès et al., 2016 ; McIntyre et al., 2011). Cette thèse conforte ces conclusions en montrant qu'un agent virtuel

adoptant des comportements de rapport virtuel, en plus d'apporter ces explications, renforce le sentiment de réussite en mathématiques. Cependant, malgré des performances similaires, les filles ont rapporté un sentiment de réussite moindre par rapport aux garçons (CATSAMBIS, 1994 ; ZANDER et al., 2020). Néanmoins, lorsque le rapport était exprimé par un agent féminin, les performances en mathématiques augmentaient tant chez les filles que chez les garçons. On peut toutefois s'interroger sur la possible contribution du sentiment d'anonymat lors des interactions avec un agent virtuel, perçu comme dépourvu de jugement humain et favorisant ainsi une plus grande ouverture de la part des individus (DEVAULT et al., 2014 ; LUCAS et al., 2017). Cette absence de jugement ressenti pourrait expliquer en partie l'amélioration des performances et du sentiment de réussite. **En résumé, si les travaux précédents ont déjà souligné l'importance d'expliquer sa réussite par le travail et les efforts, les résultats de cette thèse ont mis en lumière le rôle crucial de la notion de rapport virtuel pour incarner un modèle de réussite auquel les individus peuvent s'identifier.**

Tout comme dans les études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016) qui ont utilisé des histoires personnelles de modèles de réussite, nos agents virtuels ont expliqué leur réussite en mathématiques par le travail et les efforts. La différence avec ces études réside dans l'aspect visuel et interactif introduit par les agents virtuels. En effet, dans les études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016), les élèves devaient lire des histoires personnelles sur des modèles de réussite. Leurs résultats ont montré que le genre du modèle n'avait pas d'importance pour améliorer les performances en mathématiques des élèves, et en particulier des filles. Cependant, nos résultats ont montré l'inverse : le genre de l'agent est important. L'agent virtuel féminin avec des comportements de rapport a permis d'obtenir les meilleurs résultats pour les filles et les garçons. Nos résultats suggèrent donc que le genre féminin joue un rôle important dans l'amélioration des performances des élèves. En effet, dans la première phase de notre étude, lorsqu'une voix rappelait les notions clés de la leçon de mathématiques, c'est la voix féminine qui a conduit tous les élèves à de meilleures performances par rapport à la voix masculine. Elle a également permis aux filles de mieux réussir dans les exercices difficiles, et aux garçons de mieux réussir dans les exercices faciles. Contrairement aux résultats de MARX et ROMAN (2002), où les hommes ont moins bien réussi en mathématiques après avoir été exposés à un modèle de réussite féminin très compétent en mathématiques, les garçons de notre étude ont mieux réussi les tests de mathématiques avec la voix féminine au temps 1, et avec l'agent féminin avec des comportements de rapport au temps 2 (tout comme les filles). Bien que nous n'ayons pas mesuré le niveau de compétences perçu en mathématiques des agents dans l'étude sur l'impact des agents virtuels sur les performances (Chapitre 9), nos deux études sur l'évaluation des agents virtuels (Chapitre 7) ont montré que tous les élèves, filles comme garçons, ont perçu les agents féminins et masculins comme étant très compétents en mathématiques. **En résumé, contrairement à certains travaux antérieurs (BAGÈS & MARTINOT, 2011 ; BAGÈS et al., 2016), nos résultats montrent que le genre de l'agent virtuel pédagogique est un facteur clé pour améliorer les performances en mathématiques. L'agent féminin est le plus bénéfique pour tous les élèves, y compris les garçons, contrairement à d'autres études où les hommes (adultes) rejettent les modèles féminins très compétents en mathématiques.**

Cette différence de résultats par rapport à ces travaux antérieurs (BAGÈS & MARTINOT, 2011 ; BAGÈS et al., 2016 ; MARX & ROMAN, 2002) peut s'expliquer par plusieurs facteurs. Tout d'abord, nous pensons que l'interaction avec l'agent virtuel a rendu leur genre plus saillant que dans

les études de BAGÈS et MARTINOT (2011) et BAGÈS et al. (2016) où il n'y avait aucun visuel. On peut se questionner sur les raisons pour lesquelles c'est le genre féminin qui a été bénéfique pour tous les élèves. Les résultats de nos études perceptives (Chapitre 7) apportent certains éclairages. Tout d'abord, les filles ont ressenti plus de rapport virtuel avec l'agent féminin qu'avec l'agent masculin. Aussi, le rapport virtuel a notamment permis aux élèves de mieux percevoir les agents comme des modèles de réussite, et de mieux s'identifier à eux. De plus, compte tenu de la littérature sur les profils des modèles de réussite pour réduire les effets de menace du stéréotype, il n'est pas étonnant que les filles aient mieux réussi avec l'agent féminin (LEROY et al., 2022). Concernant les garçons, une possible interprétation repose sur le lien entre la féminité, le sentiment de rapport, et le niveau d'anthropomorphisme. En effet, une récente étude a montré que les femmes ainsi que les robots féminins sont perçus comme plus humains que leur homologues masculins (BORAU et al., 2021). Dans notre étude perceptive, l'agent féminin avec des comportements de rapport virtuel a été jugé plus proche de l'humain que ce même agent féminin sans comportements de rapport virtuel. Dans l'étude principale, peut-être qu'un sentiment de rapport plus fort a été perçu avec l'agent féminin avec des comportements de rapport virtuel, ce qui a entrainé une plus grande perception d'anthropomorphisme. Ces trois facteurs (anthropomorphisme, sentiment de rapport, genre féminin) pourraient expliquer pourquoi l'agent féminin avec des comportements de rapport a également permis aux garçons d'obtenir les meilleurs résultats au test standardisé de mathématiques. Une autre interprétation possible est que les garçons auraient pu situer leur réussite en mathématiques par rapport à celle de l'agent féminin, pensant que si elle peut réussir, alors eux aussi, ce serait a priori facile. Cependant, les résultats des études perceptives ont montré que les garçons trouvaient l'agent féminin très compétent en mathématiques.

10.3. Implications pratiques

Notre travail offre également des contributions pratiques importantes, notamment en créant un environnement virtuel d'apprentissage sain qui protège l'identité de genre des jeunes filles. En effet, une façon de créer un environnement favorable pour les filles en mathématiques est de mettre en valeur leur identité de genre, en montrant que cela ne constitue pas un frein à leur réussite dans cette matière. Cela implique de réaffirmer les compétences des filles, notamment en mathématiques, en mettant en avant notre agent virtuel pédagogique féminin compétent dans ce domaine et manifestant des comportements de rapport virtuel. Bien que le numérique éducatif soit très présent dans les classes, il est rarement évalué. Dans cette thèse, nous avons montré qu'un agent virtuel féminin, créé à partir des travaux sur la menace du stéréotype et incarnant un modèle de réussite avec un comportement de rapport virtuel, est bénéfique pour l'enseignement des mathématiques. Ce type d'agent virtuel pédagogique pourrait être utilisé à la fois pour les garçons et les filles, que ce soit pour renforcer les connaissances en classe ou à la maison. Pour les enseignants, il peut constituer un outil précieux pour travailler en petits groupes, certains s'appuyant sur des agents pédagogiques et d'autres bénéficiant de l'interaction directe avec l'enseignant.

L'utilisation régulière de notre agent virtuel féminin pourrait permettre de renforcer la confiance des filles dans leur capacité de réussite en mathématiques. En effet, les résultats de nos études montrent que cet agent contribue à accroître le sentiment de réussite chez les

filles, tout en potentiellement atténuant les stéréotypes de genre. En outre, une exposition fréquente des garçons à la réussite mathématique d'une entité virtuelle féminine pourrait aider à remettre en question les stéréotypes traditionnels. Les travaux montrant l'effet persistant dans le monde réel des expériences dans le monde virtuel (OCHS, 2020 ; VICENTE & MATUTE, 2023 ; WEST et al., 2019) nous amènent à penser que l'exposition prolongée à des modèles de réussite virtuels féminins pourrait contribuer à réduire les stéréotypes de genre dans le domaine des sciences.

Nos travaux fournissent également des informations intéressantes sur l'effet des systèmes de notations sur les performances et la stéréotypie implicite des élèves. La notation par couleur (e.g., attribuer à l'élève un groupe de compétences plutôt qu'une note) semble diminuer, voire éliminer, les disparités entre filles et garçons en mathématiques. En effet, les élèves avec ce système de notation présentaient moins de stéréotype implicite, notamment chez les filles où l'absence de stéréotypie implicite, voire une contre-stéréotypie, était observée. Néanmoins, l'utilisation de la notation par couleur était associée à des performances plus faibles en mathématiques par rapport aux élèves notés sur une échelle classique de 0 à 20. Ainsi, pour réduire la stéréotypie implicite tout en préservant de bonnes performances, une approche consisterait à combiner l'utilisation de la notation classique avec une exposition régulière des élèves à des agents virtuels pédagogiques féminins, comme suggéré dans notre thèse. Étant donné que nos résultats montrent que la stéréotypie implicite reste malléable à cet âge, une exposition prolongée de nos participants à notre agent virtuel féminin incarnant un modèle de réussite pourrait potentiellement modifier les automatismes de stéréotypie implicite. Dans cette perspective, les élèves pourraient associer aussi bien les garçons que les filles aux mathématiques, tout en améliorant leurs performances dans cette matière. Il est à noter également que nos agents avec des comportements de rapport virtuel, peu importe leur genre, ont amélioré le sentiment de réussite en mathématiques des élèves : une exposition continue pourrait alors aider à renforcer le sentiment de réussite des élèves dans cette discipline.

10.4. Implications éthiques

Nous mettons en relation nos travaux de thèse avec les travaux sur l'éthique en intelligence artificielle, et plus particulièrement dans le cadre des agents virtuels. SCHLIMBACH et KHOSRAWI-RAD (2022) ont proposé une séries de recommandations sur les dimensions éthiques à considérer lors de la conception d'agents virtuels pédagogiques (e.g., apparence, comportements) en s'appuyant sur les cinq grands principes éthiques de l'OCDE (2019) (p. 1-2) sur l'intelligence artificielle :

1. "L'intelligence artificielle devrait profiter aux individus et à la planète en favorisant la croissance inclusive, le développement durable et l'amélioration du bien-être."
2. "Les systèmes d'intelligence artificielle devraient être conçus dans une optique de respect de l'état de droit, des droits de l'homme, des valeurs démocratiques et de la diversité. Ils devraient présenter des garanties appropriées – par exemple, en permettant une intervention humaine lorsque nécessaire – afin de tendre vers une société équitable et juste."

3. "Il convient de veiller à la transparence et à une divulgation responsable des informations liées aux systèmes d'intelligence artificielle afin de s'assurer que les individus puissent en comprendre et en contester les résultats."
4. "Les systèmes d'intelligence artificielle devraient être robustes, sûrs et sécurisés tout au long de leur cycle de vie et les risques potentiels devraient être évalués et gérés de manière continue."
5. "Les organisations et les individus concevant, déployant ou exploitant des systèmes d'intelligence artificielle devraient être tenus responsables de leur bon fonctionnement, dans le respect des principes exposés ci-dessus."

Les recommandations proposés par SCHLIMBACH et KHOSRAWI-RAD (2022) relativement aux 5 principes ci-dessus portent sur plusieurs aspects des agents virtuels. Nous discutons ces recommandations plus en détail dans cette section, plus particulièrement concernant l'apparence de nos agents virtuels pédagogiques, puis de leurs comportements verbaux et non-verbaux.

10.4.1. Apparence des agents virtuels pédagogiques

L'apparence des agents virtuels pédagogiques peut être reliée au deuxième principe de l'éthique en intelligence artificielle de l'OCDE (2019) : les droits humains et l'équité. Ce principe se concentre plus particulièrement sur les discriminations. L'étude de SCHLIMBACH et KHOSRAWI-RAD (2022) s'est penchée sur l'aspect des agents virtuels pédagogiques en vue de représenter le plus de personnes possibles, pour ne pas systématiquement créer les mêmes agents. Deux méthodes y sont décrites : créer des agents non-humains (comme des animaux ou des robots), ou laisser l'utilisateur créer l'agent virtuel lui-même et le laisser ainsi choisir le genre, les vêtements, la coupe de cheveux, et d'autres caractéristiques. Nous n'avons pas suivi ces recommandations dans nos recherches pour plusieurs raisons.

Tout d'abord, la première méthode ne lutte ni contre les stéréotypes de genre, ni contre les effets de menace du stéréotype. Dans le domaine de l'éducatif numérique, la volonté qui se cache souvent derrière le choix des animaux ou robots virtuels est de ne pas associer un genre ou une origine culturelle à un agent virtuel, de sorte à ce qu'aucun apprenant ne se sente délaissé. Cependant, des recherches ayant exploré l'utilisation d'agents virtuels sans genre (e.g., mannequin en bois) ont montré que les utilisateurs attribuent tout de même un genre binaire, notamment selon les mouvements de ces agents (MCDONNELL et al., 2009 ; ZIBREK et al., 2020). Comme énoncé au Chapitre 3, nous nous demandons si l'attribution du genre dépend du contexte ou du rôle des agents virtuels. Les domaines des STIMs étant considérés comme masculins (MAKAROVA et al., 2019), les utilisateurs pourraient attribuer le genre *masculin* à ce type d'agents virtuels (e.g., robots, animaux). Cela pourrait renforcer le stéréotype selon lequel les domaines des STIMs sont plus adaptés aux hommes qu'aux femmes. En ce qui concerne les rôles des agents, il y a plus d'assistantes virtuelles que d'assistants virtuels, souvent à cause de stéréotypes de genre car les femmes sont censées servir, aider et entretenir les autres (BRAHNAM & WEAVER, 2015) ; bien que d'autres études aient montré que les robots à l'apparence féminine et les assistants vocaux à la voix féminine sont considérés comme plus humains (BORAU et al., 2021). Cependant, les assistants virtuels agenres pourraient être considérés comme des femmes, et renforcer les stéréotypes néfastes sur les

femmes (WEST et al., 2019). Enfin, une étude de MAKRANSKY et al. (2019) a montré qu'un agent féminin représentant une scientifique (e.g., en portant une blouse blanche dans un contexte de science) permettait à de jeunes apprenantes de mieux apprendre et mieux réussir un test sur la sécurité en chimie, comparativement à un agent virtuel ayant l'apparence d'un drone. Les agents ou robots à l'apparence d'animaux peuvent tout de même être bénéfiques, par exemple en contexte médical. On peut par exemple citer Paro, une peluche robotisée en forme de bébé phoque destinée aux personnes âgées pour diminuer le stress (AMINUDDIN et al., 2016).

La deuxième méthode se base sur le choix des utilisateurs pour créer leur propre agent virtuel pédagogique. L'idée est que les élèves pourraient ainsi créer un modèle de réussite adapté pour eux-mêmes. SCHLIMBACH et KHOSRAWI-RAD (2022) apportent une limite à cela : les utilisateurs peuvent ne pas savoir ce qui serait la meilleure représentation pour eux. Les élèves pourraient penser qu'un agent virtuel pédagogique masculin à l'aspect d'adulte correspondrait le mieux pour les soutenir en mathématiques, en accord avec les stéréotypes de genre. Nous avons montré au Chapitre 2 à quel point les stéréotypes de genre pouvaient influencer les enfants. Des études montrent notamment que les enfants ont tendance à dessiner des hommes lorsqu'on leur demande de représenter une personne scientifique.

Compte tenu de nos résultats sur l'impact positif de l'agent féminin sur les performances en mathématiques, et dans l'optique d'encourager le plus de personnes à améliorer leurs performances dans un domaine et à se sentir représenté, nous soutenons l'idée de créer des agents virtuels pédagogiques contre-stéréotypés (e.g., des agents féminins en sciences, des agents masculins en lecture ou en tant qu'assistant) pour d'une part ne pas associer un genre à une activité ou un rôle, et d'autre part pour lutter contre la menace du stéréotype dans divers domaines. Les résultats de nos études vont dans ce sens : l'agent virtuel pédagogique féminin a permis aux filles et aux garçons d'obtenir les meilleurs résultats en mathématiques. D'autres recherches ont montré qu'un agent virtuel pédagogique masculin a permis aux élèves d'obtenir de meilleurs résultats en lecture (Y. KIM, 2013), un sujet où l'écart de performance entre filles et garçons (cette fois-ci en faveur des filles) est encore plus grand que celui en mathématiques en faveur des garçons. Cela laisse cependant de côté d'autres caractéristiques liées à l'apparence, pourtant tout aussi importants, comme l'origine culturelle. Nous avons fait le choix de prendre en considération l'origine culturelle prédominante des élèves participants à nos études. Nous avons donné à nos agents l'apparence d'une jeune fille maghrébine et d'un jeune garçon maghrébin (couleur de peau plus ou moins foncée, cheveux non lisses, couleurs des yeux foncés), bien que cela ne représente pas les élèves ayant d'autres origines.

10.4.2. Comportements verbaux et non-verbaux des agents virtuels pédagogiques

Concernant le comportements de nos agents, nous avons créé un modèle de rapport virtuel à partir de la littérature sur le rapport dans les relations interpersonnelles, mais également dans les interactions humain-agent. Notre modèle se fonde sur les gestes, les postures et les expressions faciales qui favorisent le rapport virtuel. Les résultats des évaluations perceptives de nos agents virtuels pédagogiques ont tout d'abord montré que l'agent féminin avec des

comportements de rapport virtuel a été jugé comme plus humain que l'agent féminin sans rapport virtuel. C'était l'inverse pour les agents masculins. Il semblerait qu'un agent féminin doit être souriant et expressif, à l'inverse d'un agent masculin qui devrait être plutôt neutre et sans expression positive, pour être considéré comme plus humain. Ceci est cohérent avec les études en sciences sociales, montrant que les femmes tendent à sourire plus que les hommes (LaFrance & Vial, 2016). De plus, l'agent féminin a été significativement perçu comme générant plus de rapport virtuel que l'agent masculin. Dans notre étude principale, l'agent féminin avec des comportements de rapport virtuel était associé aux meilleurs résultats au test standardisé de mathématiques. En se fondant sur les résultats de nos études pilotes perceptives, on peut supposer qu'un sentiment de rapport plus fort ait été ressenti avec tous les agents féminins lors de notre étude principale, et encore plus avec l'agent féminin exprimant des comportements de rapport virtuel. Cet agent aurait été perçu comme plus proche de l'humain que les autres. Un niveau d'anthropomorphisme plus élevé, résultant de l'interaction entre le genre féminin de l'agent et ses comportements de rapport virtuel, pourrait expliquer pourquoi l'agent féminin avec des comportements de rapport virtuel a permis à tous les élèves d'obtenir les meilleurs résultats au test standardisé de mathématiques. Cela expliquerait également pourquoi, contrairement aux études de Bagès et Martinot (2011) et Bagès et al. (2016) où les élèves devaient lire des histoires sur un modèle de réussite sans interagir avec, le genre du modèle de réussite a une importance.

Nos résultats suggèrent donc d'utiliser un agent virtuel féminin, incarnant un modèle de réussite, avec des comportements de rapport virtuel. Cependant, lier systématiquement les comportements de rapport virtuel aux agents féminins pourraient poser des problèmes éthiques. En voulant améliorer les performances en mathématiques des élèves en présentant ce genre d'agents régulièrement, il ne faudrait pas renforcer d'autres stéréotypes de genre (e.g., les filles et les femmes devraient toujours sourire et être positives, Deutsch et al., 1987 ; LaFrance et Hecht, 1995). Il est donc important d'avoir une attention particulière sur le comportement de l'agent pour éviter de renforcer les stéréotypes de femmes dociles et serviables (West et al., 2019). Les agents virtuels féminins sont en effet significativement exposés à plus d'insultes que les agents virtuels masculins (Brahnam & De Angeli, 2012 ; De Angeli & Brahnam, 2006). Ces problématiques éthiques sont au cœur de l'actualité sur la question du genre des agents virtuels, en particulier des assistants vocaux où certaines assistantes virtuelles connues comme Siri répondaient favorablement aux insultes (West et al., 2019). Les agents virtuels ne doivent donc pas toujours adopter un comportement de rapport virtuel au risque de renforcer un stéréotype existant sur le comportement attendu des femmes dans le monde réel. Des travaux montrent en effet que les stéréotypes et autres biais rencontrés dans le monde virtuel persistent dans le monde réel (Ochs, 2020 ; Vicente & Matute, 2023 ; West et al., 2019).

Bien que nos résultats indiquent que le rapport virtuel favorise l'identification à l'agent et améliore les résultats en mathématiques des élèves, l'utilisation d'agents dotés de comportements de rapport soulève des questions éthiques liées aux risques de manipulation des individus. Une caractéristique spécifique des interactions avec ces agents virtuels est le *sentiment d'anonymat*, comme l'ont souligné Lugrin et al. (2022). Le sentiment d'anonymat peut encourager des effets positifs du rapport, comme révéler plus d'informations sur soi-même. Des recherches montrent en effet qu'un plus grand sentiment de rapport permet aux personnes de plus se confier (Gratch et al., 2007). Les agents virtuels, en tant que pro-

grammes informatiques a priori non contrôlés par un humain, permettent aux individus de ressentir un certain anonymat leur permettant de se sentir plus à l'aise et de révéler des informations personnelles, sans crainte d'être jugés (DEVAULT et al., 2014) ou discriminés (VILARO et al., 2021). Le sentiment d'anonymat permet également de préserver l'identité de l'utilisateur. Cependant, cela pose des questions éthiques. On pourrait en effet utiliser le rapport virtuel et le sentiment d'anonymat pour inciter les individus à révéler des données sensibles. Cela fait référence au troisième principe de l'OCDE (2019) sur la transparence et la responsabilité des systèmes d'intelligence artificielle. SCHLIMBACH et KHOSRAWI-RAD (2022) ont enrichi ce principe avec les règles suivantes : les agents virtuels doivent démontrer comment les données des utilisateurs sont analysées, se conformer aux Règlement Général sur la Protection des Données (RGPD), et ne pas abuser de la confiance qu'un utilisateur pourrait leur accorder. Modéliser le rapport virtuel reste à ce jour un défi dans le domaine des agents virtuels, que ce soit du côté éthique ou du côté computationnel (CLAVEL et al., 2022).

Enfin, le premier principe de l'éthique pour les systèmes d'intelligence artificielle de l'OCDE (2019) est la durabilité et le bien-être des utilisateurs. Des risques d'attachement à l'application, plus particulièrement à l'agent virtuel pédagogique avec le modèle de rapport, pourraient s'observer. Pour éviter une utilisation excessive de l'application, SCHLIMBACH et KHOSRAWI-RAD (2022) recommandent d'intégrer des messages pour encourager les utilisateurs à prendre une pause, par exemple avec des exercices de relaxation, ou encore de fixer des périodes d'utilisation.

10.4.3. Quelques réflexions personnelles

Mes réflexions personnelles portent plus particulièrement sur l'apparence de nos agents virtuels pédagogiques, et plus généralement sur l'apparence et les loisirs des modèles de réussite incarnés par des femmes (virtuels ou non). Quelques études ont montré que la féminité pouvait entraver l'efficacité de ces modèles de réussite en mathématiques, car la combinaison de la féminité et de la réussite dans les STIMs semble inatteignable pour les jeunes filles (BETZ & SEKAQUAPTEWA, 2012 ; PRONIN et al., 2004). Ainsi, porter des vêtements roses, aimer lire des magazines (BETZ & SEKAQUAPTEWA, 2012), ou encore porter du maquillage (PRONIN et al., 2004) semble incompatible avec la réussite en mathématiques. De même, un modèle "trop" stéréotypé par rapport aux sciences (e.g., loisirs et apparence "geeks") ne permet pas aux jeunes filles de s'identifier à un modèle de réussite incarné par une femme. Finalement, c'est une apparence et des loisirs plus "neutres" (e.g., porter des vêtements foncés, aimer lire) qui est plutôt préconisé en suivant les résultats de ces études. Cela soulève encore des problématiques éthiques : si l'on n'expose jamais de modèle de réussite "trop" féminin ou "trop" masculin, nous isolons une partie des personnes en réussite en mathématiques, déjà peu représentées par des femmes. De plus, nous risquons d'aller dans le sens des stéréotypes de genre, à savoir qu'une femme féminine ne pourrait quasiment jamais réussir en mathématiques. Aussi, on peut constater que le "neutre" en ce qui concerne les vêtements rentre plutôt dans une notion de masculinité (e.g., pantalon, couleurs foncés), mais jamais dans une notion de féminité. La neutralité des vêtements n'est jamais représentée par des vêtements plutôt associés aux femmes. Cela efface les femmes considérées comme féminines ou "trop" masculines, et renvoie le message que pour réussir en sciences, les jeunes filles dans ces catégories devraient abandonner une partie de leur identité.

10.5. Limites

Comme dans toute étude réalisée sur le terrain, des biais inhérents à l'expérimentation en situation de classe quasi-ordinaire étaient présents dans nos études. D'autres limites, certaines pouvant constituer des perspectives d'amélioration de nos travaux, sont également constatées et discutées dans cette section. Nous discutons tout d'abord des limites de l'expérimentation, puis des limites de nos agents virtuels pédagogiques.

10.5.1. Limites des expérimentations

L'absence de différence de performance entre les filles et les garçons peut être attribuée en partie à la composition majoritairement en REP+ des classes participantes à l'étude. Cette décision s'est fondée sur la nécessité de tester les agents virtuels pédagogiques auprès d'élèves issus de milieux moins favorisés, qui pourraient bénéficier de dispositifs de soutien dans les classes. Néanmoins, cette approche comporte un risque car il est généralement plus facile de détecter les effets de menace du stéréotype chez une population plus favorisée. De plus, les enseignants des classes participantes avaient deux méthodes différentes pour noter leurs élèves en classe. Certains utilisaient une notation traditionnelle sur une échelle de 0 à 20, comme il est le plus courant en France dans les collèges. D'autres enseignants affectaient les élèves à l'aide d'une couleur dans un groupe de compétence : les élèves n'avaient donc pas de moyenne en mathématiques, mais une couleur. Cette différence de pratiques de notation dans les classes a sûrement introduit des biais. La notation par couleur a pu empêcher de faire ressortir les effets de menace du stéréotype, en masquant les différences de performance entre les filles et les garçons. Cela ne signifie pas nécessairement que la menace du stéréotype n'était pas présente, mais simplement qu'il est plus difficile de l'observer.

Dans notre étude principale, nous avons obtenu un plus petit nombre de données exploitables (375 au temps 2) comparé au nombre d'élèves inscrits pour participer (574 élèves) à cause des absences, des tricheries, ou encore des refus de participation. Les classes d'un même établissement et d'un même niveau (e.g., 6e ou 5e) ont effectué les passations en parallèles pour éviter que les élèves ayant déjà passé l'étude ne divulguent des informations détaillées à leurs camarades d'autres classes. L'expérimentatrice a donc pris part à certaines passations des évaluations dans leur intégralité, laissant 8 classes sur 23 être surveillées uniquement par un enseignant (pas forcément de mathématiques), parfois par un assistant éducatif. Par conséquent, certains biais ont pu être présents comme l'introduction de consignes non-sollicitées ou des libertés prises avec le protocole). L'expérimentatrice était présente dans chaque établissement lors des passations, et a pu discuter avec le personnel en charge de la surveillance des classes pour avoir un retour de terrain. Néanmoins, il n'est pas certain que toutes les classes, y compris le personnel en charge de la surveillance, aient participé sérieusement à l'étude. De plus, le fait que l'expérimentatrice se soit présentée dans chaque classe pour expliquer les consignes, même dans les classes où elle n'était pas responsable de la surveillance complète, a inévitablement introduit un biais.. Des études ont montré que le genre de l'expérimentateur pouvait avoir un effet sur la menace du stéréotype, où les filles pouvaient voir l'expérimentatrice comme un modèle de réussite, ce qui améliorait leurs résultats comparativement aux filles dont les consignes ont été données par un expérimentateur. De plus, nos expérimentations ont eu lieu uniquement en France, et plus particulièrement

en région Provence-Alpes-Côte d'Azur, ce qui ne reflète pas la diversité culturelle dans son ensemble.

Enfin, il est important de souligner que, au-delà des performances, nous ne pouvons déterminer si l'efficacité des agents résidait dans leur capacité à clarifier une notion que l'élève aurait mieux compris, ou dans leur capacité à rappeler à l'élève une notion déjà assimilée. Aussi, nos résultats sont peut-être liés à nos agents virtuels pédagogiques en particulier. Nous n'avons pas évalué d'agents avec d'autres apparences physiques. Il serait intéressant de varier les vêtements pour analyser l'impact du niveau de féminité/masculinité perçu sur les performances en mathématiques, ou encore de varier les couleurs de peau. Il aurait également été intéressant d'évaluer le niveau de féminité/masculinité perçu de nos agents virtuels. Au vu des résultats, il aurait été également intéressant d'évaluer l'identification à l'agent dans l'étude principale, et pas uniquement dans les études pilotes perceptives. Nous ne l'avons pas fait pour que les élèves aient le temps de compléter toutes les tâches en une séance d'une heure.

10.5.2. Limites des agents virtuels pédagogiques

Une limite concernant les agents virtuels pédagogiques est l'absence de variabilité dans les comportements verbaux et non-verbaux dû aux scripts. Le modèle proposé était suffisant pour les études menées dans cette thèse. Pour une utilisation plus régulière du système, la variabilité des comportements est importante. Le modèle peut cependant être étendu pour intégrer plus de variabilité comportementale, ou combiné à des approches de génération automatique de comportement comme proposé par DELBOSC et al. (2022).

Une autre limite due au contexte expérimental est le fait que nos agents ne prenaient pas en compte les signaux des élèves (e.g., expressions faciales, direction du regard, ...) pour adapter leurs comportements (e.g., demander si tout va bien si l'élève regarde ailleurs). Cela n'a pas permis de simuler l'alignement de comportements (e.g., sourire quand l'utilisateur sourit). Pour modéliser le rapport, cette composante est importante. La prise en compte des signaux de l'utilisateur nous auraient donc permis de renforcer le rapport virtuel exprimé par l'agent. Cependant, le contexte de l'étude en classe ne nous a pas permis de mettre en place la détection automatique en temps réel des signaux des élèves. Il s'agit en effet d'un environnement bruité, limité en termes d'outils de reconnaissance (pas de webcam sur les ordinateurs des collèges) et difficilement contrôlable (pertes des signaux utilisateurs si l'élève n'est pas bien en face de la webcam). De plus, le système actuel aurait pu être amélioré par un dialogue en langage naturel à partir d'un système de reconnaissance vocale. Cela aurait permis d'améliorer l'interaction avec l'agent. Cependant, étant donné encore une fois l'environnement bruité des classes contenant près d'une trentaine d'élèves, nous avons privilégié l'utilisation d'un tchat textuel pour éviter les erreurs de reconnaissance vocale qui aurait pu détériorer significativement l'interaction.

10.6. Perspectives

Malgré les limites précédemment énoncées, nos résultats ont apporté des contributions théoriques à la littérature en sciences sociales et en informatique, ainsi que des implications pratiques pour l'utilisation de nos agents en classe, notamment par la création d'un environnement virtuel d'apprentissage sécurisant pour l'identité des filles. Il serait intéressant de tester l'exposition continue à notre agent virtuel féminin avec des comportements de rapport sur les automatismes de stéréotypie implicite. Les élèves pourraient associer à la fois les garçons et les filles aux mathématiques, tout en améliorant leurs performances en mathématiques. Nous avons en effet montré qu'à cet âge-là, la stéréotypie implicite peut être encore malléable.

Concernant l'apparence de nos agents virtuels pédagogiques, nos résultats sur le niveau d'anthropomorphisme perçu montrent que les élèves n'avaient pas de doute au fait que les agents étaient bien virtuels. Les graphismes pourraient constituer un enjeu majeur dans le domaine de l'éducation numérique, à condition de ne pas tomber dans le phénomène de la vallée de l'étrange, où un décalage entre la perception de l'humanité accordée à l'apparence de l'agent virtuel et à celle accordée à ses mouvements peut engendrer un sentiment de malaise et un rejet de l'agent virtuel. Une perspective de ces recherche serait l'étude de l'impact du niveau de réalisme sur les performances en mathématiques des élèves. Pour aller plus loin, l'utilisation de la réalité virtuelle peut augmenter la présence sociale et l'engagement des apprenants. Puisque nous avons implémenté notre système sur Unity, le portage en réalité virtuelle peut être facilement réalisable. Les casques de réalité actuels (Quest Pro) permettraient de plus de détecter automatiquement les signaux des apprenants (expressions faciales, direction du regard) et permettraient ainsi d'enrichir le modèle compotuationnel de rapport proposé grâce à une capacité d'alignement des signaux de l'agent sur ceux de l'élève.

En effet, dans l'article de SCHLIMBACH et KHOSRAWI-RAD (2022), les auteurs préconisent l'adaptation du comportement verbal et non-verbal de l'agent virtuel en fonction de l'état émotionnel de l'apprenant (e.g., anxiété à propos du test, motivation à apprendre, frustration due au manque de défis). La détection des signaux de l'élève ouvrent des perspectives de reconnaissance d'états socio-affectifs tels que l'ennui, la frustration ou la joie. Le comportement de l'agent pourrait alors s'adapter à ces états reconnus (e.g., difficulté des exercices, encouragements) pour amener l'élève à un état socio-affectif optimal pour l'apprentissage. Une perspective d'amélioration de nos agents virtuels pédagogiques réside dans la prise en compte des signaux des utilisateurs pour percevoir l'état émotionnel de l'élève à l'aide de micro, webcam, et outils open-source (OpenFace, BALTRUSAITIS et al., 2018, OpenSmile, EYBEN et al., 2010) pour analyser en temps réel le comportement verbal et non-verbal des utilisateurs. Cela permettrait ainsi d'adapter les comportements de l'agent à ceux de l'élève. Des recherches utilisent ces signaux pour estimer un niveau de rapport pendant l'interaction avec les agents virtuels, afin qu'ils puissent s'adapter pour augmenter ou maintenir ce niveau de rapport. Au vu de la nature de nos expérimentations (toutes directement en classe, sans passation individuelle), il n'était pas envisageable pour nous d'utiliser ce genre de méthode à cause de la forte probabilité de ne pas détecter ces signaux. Il faut en effet se tenir droit, bien en face de la caméra pour que le visage soit reconnu. Il serait cependant possible de créer ce genre de modèle computationnel pour une utilisation de l'application à la maison. Analyser les postures du corps ou les expressions faciales permettrait d'augmenter le niveau de coordination ressenti entre l'agent et l'utilisateur, et ainsi améliorer le niveau de rapport

virtuel. Cependant, le texte de loi sur l'intelligence artificielle tel que proposé par la Commission Européenne en janvier 2024 (*The AI Act*, COUNCIL OF THE EUROPEAN COMMISSION, 2024) considère l'interdiction des systèmes de reconnaissance émotionnelle visant à déceler ou inférer les émotions ou intentions des utilisateurs à partir de leurs données biométriques, notamment dans le domaine de l'éducation. Cette interdiction fait référence aux émotions telles que le bonheur, la tristesse, la colère, la surprise, le dégoût, l'embarras, l'excitation, la honte, le mépris, la satisfaction et l'amusement. Elle exclut cependant les états physiques comme la douleur ou la fatigue (e.g., les systèmes détectant la fatigue des pilotes ou des conducteurs professionnels pour prévenir les accidents). Bien que cette directive ne concerne pas la simple reconnaissance des expressions, des gestes, ou autres mouvements, il serait interdit d'utiliser ces données pour déduire ou identifier des émotions. Cela inclus des expressions faciales basiques telles qu'un froncement de sourcils ou un sourire, des gestes comme le mouvement des mains, des bras ou de la tête, ou encore des caractéristiques vocales comme une voix élevée ou un chuchotement. Cette interdiction ne concernerait pas les chercheurs, et serait en vigueur uniquement pour les industries. Cependant, il serait impossible de proposer ce genre d'agents virtuels dans les classes ou à la maison, en dehors des recherches.

En vue d'une utilisation régulière dans les classes ou à la maison, il serait intéressant de généraliser l'outil à d'autres leçons et exercices de mathématiques. Ces leçons et exercices pourraient être générées automatiques avec des LLMs (*Large Langage Models*), pouvant être ensuite ajustés par les enseignants.

Enfin, une perspective intéressante et nécessaire serait de tester si l'agent féminin avec des comportements de rapport est également efficace dans les classes non-REP, et pas uniquement dans les classes REP+, pour savoir s'il est aussi efficace sur la performance en mathématiques des garçons qui ne sont pas en décrochage scolaire. Cela nous permettrait de trancher sur la question de l'efficacité de cet agent sur la menace du stéréotype, et d'obtenir davantage de réponses à nos interrogations. En effet, comme expliqué plus haut, une interprétation de nos résultats est que cet agent féminin était bénéfique pour les filles par réduction de menace du stéréotype, et pour les garçons en décrochage scolaire en leur apportant un soutien nécessaire. Une question se pose alors : les élèves les plus au cœur des stéréotypes, c'est-à-dire les garçons fort en mathématiques dans des environnements plus favorisés, rejetteraient-ils cet agent, comme les hommes adultes ont pu rejeter les modèles de réussite féminin très compétents en mathématiques (MARX & ROMAN, 2002) ?

Bibliographie

AIT CHALLAL, T., & GRYNSZPAN, O. (2018). What Gaze Tells Us About Personality. *Proceedings of the 6th International Conference on Human-Agent Interaction*, 129-137.

AKBAR, F., GROVER, T., MARK, G., & ZHOU, M. X. (2018). The Effects of Virtual Agents' Characteristics on User Impressions and Language Use. *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, 1-2.

AMBADY, N., PAIK, S. K., STEELE, J., OWEN-SMITH, A., & MITCHELL, J. P. (2004). Deflecting negative self-relevant stereotype activation : The effects of individuation. *Journal of Experimental Social Psychology*, *40*(3), 401-408.

AMBADY, N., & ROSENTHAL, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences : A meta-analysis. *Psychological bulletin*, *111*(2), 256.

AMBADY, N., SHIH, M., KIM, A., & PITTINSKY, T. L. (2001). Stereotype susceptibility in children : Effects of identity activation on quantitative performance. *Psychological science*, *12*(5), 385-390.

AMINUDDIN, R., SHARKEY, A., & LEVITA, L. (2016). Interaction with the Paro robot may reduce psychophysiological stress responses. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 593-594.

ANDRE, T., WHIGHAM, M., HENDRICKSON, A., & CHAMBERS, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching : The Official Journal of the National Association for Research in Science Teaching*, *36*(6), 719-747.

ANDREU, S., BALLEREAU, M.-A., ALI, L., BRET, A., SANTOS, R., DURAND DE MONESTROL, H., HEIDMANN, L., LAMBERT, K., M'BAFOUMOU, A., PAILLET, V., PAUL, A., PERSEM, É., ROCHER, T., RUE, G., STACHOWIAK, J.-F., SALLES, F., & VOURC', R. (2022). Évaluations de début de sixième en 2021 [Chaque note d'information présente les résultats les plus récents issus des exploitations d'enquêtes et d'études statistiques. Cette publication apporte l'éclairage de la direction de l'évaluation, de la prospective et de la performance (DEPP) à la compréhension et à l'analyse du système éducatif français. Elle est disponible en ligne, sous la forme de fichiers imprimables à télécharger, accompagnés de tableaux de données chiffrées et de graphiques.]. https://doi.org/10.48464/ni-22-04

ANDREU, S., BALLEREAU, M.-A., BRET, A., CHAAYA, C., DURAND DE MONESTROL, H., HEIDMANN, L., M'BAFOUMOU, A., PAILLET, V., PAUL, A., PERSEM, E., ROCHER, T., RUE, G., SATHICQ, V., STACHOWIAK, J.-F., SALLES, F., & VOURC', R. (2023). Évaluation exhaustive de début de sixième 2022 (d. l. p. e. d. l. p. MINISTÈRE

DE L'EDUCATION NATIONALE ET DE LA JEUNESSE. DIRECTION DE L'ÉVALUATION, Éd.) [Chaque note d'information présente les résultats les plus récents issus des exploitations d'enquêtes et d'études statistiques. Cette publication apporte l'éclairage de la direction de l'évaluation, de la prospective et de la performance (DEPP) à la compréhension et à l'analyse du système éducatif français. Elle est disponible en ligne, sous la forme de fichiers imprimables à télécharger, accompagnés de tableaux de données chiffrées et de graphiques.]. https://doi.org/10.48464/ni-23-03

ANTONIO GÓMEZ-JÁUREGUI, D., GIRAUD, T., ISABLEU, B., & MARTIN, J.-C. (2021). Design and evaluation of postural interactions between users and a listening virtual agent during a simulated job interview. *Computer Animation and Virtual Worlds*, e2029.

ARGYLE, M., & DEAN, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289-304.

ARMANDO, M., OCHS, M., & RÉGNER, I. (2022). The Impact of Pedagogical Agents' Gender on Academic Learning : A Systematic Review. *Frontiers in Artificial Intelligence*, *5*.

ARMANDO, M., RÉGNER, I., & OCHS, M. (2021). Les effets du genre des agents virtuels sur l'apprentissage. *WACAI 2021*.

ARMANDO, M., RÉGNER, I., & OCHS, M. (2023). Toward a Tool Against Stereotype Threat in Math : Children's Perceptions of Virtual Role Models. In *International Conference on Multimodal Interaction* (p. 306-310).

ARNOLD, E. G., BURROUGHS, E. A., & DESHLER, J. M. (2020). Investigating classroom implementation of research-based interventions for reducing stereotype threat in calculus. *International Journal of Research & Method in Education*, *43*(1), 67-77.

ARONSON, J., FRIED, C. B., & GOOD, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of experimental social psychology*, *38*(2), 113-125.

ARONSON, J., LUSTINA, M. J., GOOD, C., KEOUGH, K., STEELE, C. M., & BROWN, J. (1999). When white men can't do math : Necessary and sufficient factors in stereotype threat. *Journal of experimental social psychology*, *35*(1), 29-46.

ASTRID, M., KRÄMER, N. C., & GRATCH, J. (2010). How our personality shapes our interactions with virtual characters-implications for research and development. *International Conference on Intelligent Virtual Agents*, 208-221.

ATKINSON, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, *94*(2), 416.

BAGÈS, C., & MARTINOT, D. (2011). What is the best model for girls and boys faced with a standardized mathematics evaluation situation : A hardworking role model or a gifted role model? *British Journal of Social Psychology*, *50*(3), 536-543.

BAGÈS, C., VERNIERS, C., & MARTINOT, D. (2016). Virtues of a hardworking role model to improve girls' mathematics performance. *Psychology of Women Quarterly*, *40*(1), 55-64.

BAILENSON, J. N., & YEE, N. (2005). Digital chameleons : Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological science*, *16*(10), 814-819.

BAKER, E. R., TISAK, M. S., & TISAK, J. (2016). What can boys and girls do ? Preschoolers' perspectives regarding gender roles across domains of behavior. *Social Psychology of Education*, *19*, 23-39.

BALTRUSAITIS, T., ZADEH, A., LIM, Y. C., & MORENCY, L.-P. (2018). Openface 2.0 : Facial behavior analysis toolkit. *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, 59-66.

BANDURA, A. (1986). Social foundations of thought and action : A social cognitive theory.

BARTNECK, C., KULIĆ, D., CROFT, E., & ZOGHBI, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*, 71-81.

BAYLOR, A. L., & CHANG, S. (2002). Pedagogical agents as scaffolds : The role of feedback timing, number of agents, and adaptive feedback. *International Conference of the Learning Sciences, Seattle, WA*, 3-4.

BAYLOR, A. L., & KIM, Y. (2004). Pedagogical agent design : The impact of agent realism, gender, ethnicity, and instructional role. *International Conference on Intelligent Tutoring Systems*, 592-603.

BAYLOR, A. L., & KIM, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, *15*(1), 95.

BEAMAN, L., DUFLO, E., PANDE, R., & TOPALOVA, P. (2012). Female leadership raises aspirations and educational attainment for girls : A policy experiment in India. *science*, *335*(6068), 582-586.

BEILOCK, S. L., GUNDERSON, E. A., RAMIREZ, G., & LEVINE, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

BEILOCK, S. L., JELLISON, W. A., RYDELL, R. J., MCCONNELL, A. R., & CARR, T. H. (2006). On the causal mechanisms of stereotype threat : can skills that don't rely heavily on working memory still be threatened ? *Personality and Social Psychology Bulletin*, *32*(8), 1059-1071.

BEILOCK, S. L., RYDELL, R. J., & MCCONNELL, A. R. (2007). Stereotype threat and working memory : mechanisms, alleviation, and spillover. *Journal of Experimental Psychology : General*, *136*(2), 256.

BEISHUIZEN, J., HOF, E., VAN PUTTEN, C., BOUWMEESTER, S., & ASSCHER, J. (2001). Students' and teachers' cognitions about good teachers. *British Journal of Educational Psychology*, *71*(2), 185-201.

BEM, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155.

BENBOW, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents : Their nature, effects, and possible causes. *Behavioral and Brain sciences*, *11*(2), 169-183.

BENBOW, C. P., & STANLEY, J. C. (1980). Sex differences in mathematical ability : Fact or artifact? *Science*, *210*(4475), 1262-1264.

BENSON, T. A., COHEN, A. L., & BUSKIST, W. (2005). Rapport : Its relation to student attitudes and behaviors toward teachers and classes. *Teaching of Psychology*, *32*(4), 237-239.

BEN-ZEEV, T., DUNCAN, S., & FORBES, C. (2005). Stereotypes and math performance.

BEN-ZEEV, T., FEIN, S., & INZLICHT, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*(2), 174-181.

BERNIERI, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal behavior*, *12*, 120-138.

BERNIGOLE, V., FERNANDEZ, A., LOI, M., & SALLES, F. (2023). PISA 2022 : la France ne fait pas exception à la baisse généralisée des performances en culture mathématique dans l'OCDE (d. l. p. e. d. l. p. MINISTÈRE DE L'EDUCATION NATIONALE ET DE LA JEUNESSE. DIRECTION DE L'ÉVALUATION, Éd.) [Chaque note d'information présente les résultats les plus récents issus des exploitations d'enquêtes et d'études statistiques. Cette publication apporte l'éclairage de la direction de l'évaluation, de la prospective et de la performance (DEPP) à la compréhension et à l'analyse du système éducatif français. Elle est disponible en ligne, sous la forme de fichiers imprimables à télécharger, accompagnés de tableaux de données chiffrées et de graphiques.]. https://doi.org/10.48464/ni-23-48

BETZ, D. E., & SEKAQUAPTEWA, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social psychological and personality science*, *3*(6), 738-746.

BIAN, L., LESLIE, S.-J., & CIMPIAN, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389-391.

BIOCCA, F. (1997). The cyborg's dilemma : Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication*, *3*(2), JCMC324.

BIOCCA, F., HARMS, C., & GREGG, J. (2001). The networked minds measure of social presence : Pilot test of the factor structure and concurrent validity. *4th Annual International Workshop on Presence, Philadelphia, PA*, 1-9.

BLACKWELL, L. S., TRZESNIEWSKI, K. H., & DWECK, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition : A longitudinal study and an intervention. *Child development*, *78*(1), 246-263.

BLAKEMORE, J. E. O., & CENTERS, R. E. (2005). Characteristics of boys' and girls' toys. *Sex roles*, *53*, 619-633.

BLASCOVICH, J., SPENCER, S. J., QUINN, D., & STEELE, C. (2001). African Americans and high blood pressure : The role of stereotype threat. *Psychological science*, *12*(3), 225-229.

BONNOT, V., & CROIZET, J.-C. (2011). Stereotype threat and stereotype endorsement : Their joint influence on women's math performance. *Revue internationale de psychologie sociale*, *24*(2), 105-120.

BORAU, S., OTTERBRING, T., LAPORTE, S., & FOSSO WAMBA, S. (2021). The most human bot : Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing, 38*(7), 1052-1068.

BOUCHER, K. L., RYDELL, R. J., VAN LOO, K. J., & RYDELL, M. T. (2012). Reducing stereotype threat in order to facilitate learning. *European Journal of Social Psychology, 42*(2), 174-179.

BRAHNAM, S., & DE ANGELI, A. (2012). Gender affordances of conversational agents. *Interacting with Computers, 24*(3), 139-153.

BRAHNAM, S., & WEAVER, M. (2015). Re/framing virtual conversational partners : A feminist critique and tentative move towards a new design paradigm. *International Conference of Design, User Experience, and Usability*, 172-183.

BREDA, T., GRENET, J., MONNET, M., & VAN EFFENTERRE, C. (2019). Role Models féminins : un levier efficace pour inciter les filles à poursuivre des études scientifiques ? *Notes IPP*, (45).

BREDA, T., GRENET, J., MONNET, M., & VAN EFFENTERRE, C. (2021). Do female role models reduce the gender gap in science ? Evidence from French high schools.

BREDA, T., SULTAN, J., & TOUITOU, P. L. (2024). Le décrochage des filles en mathématiques dès le CP : une dynamique diffuse dans la société.

BRODISH, A. B., & DEVINE, P. G. (2009). The role of performance–avoidance goals and worry in mediating the relationship between stereotype threat and performance. *Journal of experimental social psychology, 45*(1), 180-185.

BROWN, R. P., & PINEL, E. C. (2003). Stigma on my mind : Individual differences in the experience of stereotype threat. *Journal of experimental social psychology, 39*(6), 626-633.

BUCK, G. A., CLARK, V. L. P., LESLIE-PELECKY, D., LU, Y., & CERDA-LIZARRAGA, P. (2008). Examining the cognitive processes used by adolescent girls and women scientists in identifying science role models : A feminist approach. *Science Education, 92*(4), 688-707.

BURGOON, J. K., & LE POIRE, B. A. (1999). Nonverbal cues and interpersonal judgments : Participant and observer perceptions of intimacy, dominance, composure, and formality. *Communications Monographs, 66*(2), 105-124.

CADINU, M., MAASS, A., LOMBARDO, M., & FRIGERIO, S. (2006). Stereotype threat : The moderating role of locus of control beliefs. *European journal of social psychology, 36*(2), 183-197.

CADINU, M. R., & ROTHBART, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of personality and social psychology, 70*(4), 661.

CAMPBELL, S. M., & COLLAER, M. L. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychology of Women Quarterly, 33*(4), 437-444.

CARR, P. B., & STEELE, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology, 45*(4), 853-859.

CASSELL, J. (2001). Embodied conversational agents : representation and intelligence in user interfaces. *AI magazine, 22*(4), 67-67.

CASSELL, J., & TVERSKY, D. (2005). The language of online intercultural community formation. *Journal of Computer-Mediated Communication, 10*(2), JCMC1027.

CATSAMBIS, S. (1994). The path to math : Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of education*, 199-215.

CAUNA, A., COLÉ, P., & TELLIER, M. (2023). Quel usage du geste pédagogique en maternelle en France ? *Colloque international en éducation.*

CHALABAEV, A., MAJOR, B., SARRAZIN, P., & CURY, F. (2012). When avoiding failure improves performance : Stereotype threat and the impact of performance goals. *Motivation and Emotion, 36*, 130-142.

CHAMBERS, D. W. (1983). Stereotypic images of the scientist : The draw-a-scientist test. *Science education, 67*(2), 255-265.

CHANG, F., LUO, M., WALTON, G., AGUILAR, L., & BAILENSON, J. (2019). Stereotype threat in virtual learning environments : Effects of avatar gender and sexist behavior on women's math learning outcomes. *Cyberpsychology, Behavior, and Social Networking, 22*(10), 634-640.

CHATEIGNIER, C., DUTRÉVIS, M., NUGIER, A., & CHEKROUN, P. (2009). French-Arab students and verbal intellectual performance : Do they really suffer from a negative intellectual stereotype ? *European Journal of Psychology of Education, 24*, 219-234.

CHERYAN, S., MELTZOFF, A. N., & KIM, S. (2011). Classrooms matter : The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education, 57*(2), 1825-1835.

CHERYAN, S., SIY, J. O., VICHAYAPAI, M., DRURY, B. J., & KIM, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM ? *Social psychological and personality science, 2*(6), 656-664.

CHOI, S., & CLARK, R. E. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of educational computing research, 34*(4), 441-466.

CLARK, R. E., & CHOI, S. (2007). The questionable benefits of pedagogical agents : Response to Veletsianos. *Journal of Educational Computing Research, 36*(4), 379-381.

CLAVEL, C., LABEAU, M., & CASSELL, J. (2022). Socio-conversational systems : Three challenges at the crossroads of fields. *Frontiers in Robotics and AI, 9*, 937825.

CLOUD-BUCKNER, J., SELLICK, M., SAINATHUNI, B., YANG, B., & GALLIMORE, J. (2009). Expression of personality through avatars : Analysis of effects of gender and race on perceptions of personality. *International Conference on Human-Computer Interaction*, 248-256.

COHEN, S. B. (2003). The essential difference : The truth about the male and female brain.

COUNCIL OF THE EUROPEAN COMMISSION. (2024). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union

legislative acts - Analysis of the final compromise text with a view to agreement [https://artificialintelligenceact.eu/fr/l-acte/].

CRISP, R. J., BACHE, L. M., & MAITNER, A. T. (2009). Dynamics of social comparison in counter-stereotypic domains : Stereotype boost, not stereotype threat, for women engineering majors. *Social Influence*, *4*(3), 171-184.

CROCKER, J., & MAJOR, B. (1989). Social stigma and self-esteem : The self-protective properties of stigma. *Psychological review*, *96*(4), 608.

CROIZET, J.-C., & CLAIRE, T. (1998). Extending the concept of stereotype threat to social class : The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*(6), 588-594.

CROSNOE, R., RIEGLE-CRUMB, C., FIELD, S., FRANK, K., & MULLER, C. (2008). Peer group contexts of girls' and boys' academic experiences. *Child development*, *79*(1), 139-155.

CUPERMAN, R., & ICKES, W. (2009). Big Five predictors of behavior and perceptions in initial dyadic interactions : Personality similarity helps extraverts and introverts, but hurts "disagreeables". *Journal of personality and social psychology*, *97*(4), 667.

CVENCEK, D., KAPUR, M., & MELTZOFF, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and instruction*, *39*, 1-10.

CVENCEK, D., MELTZOFF, A. N., & GREENWALD, A. G. (2011). Math–gender stereotypes in elementary school children. *Child development*, *82*(3), 766-779.

DANAHER, K., & CRANDALL, C. S. (2008). Stereotype threat in applied settings reexamined. *Journal of Applied Social Psychology*, *38*(6), 1639-1655.

DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

DANSO, H. A., & ESSES, V. M. (2001). Black experimenters and the intellectual test performance of White participants : The tables are turned. *Journal of Experimental Social Psychology*, *37*(2), 158-165.

DASGUPTA, N., & ASGARI, S. (2004). Seeing is believing : Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of experimental social psychology*, *40*(5), 642-658.

DE ANGELI, A., & BRAHNAM, S. (2006). Sex stereotypes and conversational agents. *Proc. of Gender and Interaction : Real and Virtual Women in a Male World, Venice, Italy*.

DELBOSC, A., OCHS, M., & AYACHE, S. (2022). Automatic facial expressions, gaze direction and head movements generation of a virtual agent. *Companion Publication of the 2022 International Conference on Multimodal Interaction*, 79-88.

DERKS, B., VAN LAAR, C., & ELLEMERS, N. (2016). The queen bee phenomenon : Why women leaders distance themselves from junior women. *The Leadership Quarterly*, *27*(3), 456-469.

DEUTSCH, F. M., LEBARON, D., & FRYER, M. M. (1987). What is in a smile? *Psychology of Women Quarterly*, *11*(3), 341-352.

DEVAULT, D., ARTSTEIN, R., BENN, G., DEY, T., FAST, E., GAINER, A., GEORGILA, K., GRATCH, J., HARTHOLT, A., LHOMMET, M., et al. (2014). SimSensei Kiosk : A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, 1061-1068.

DILL, K. E., BROWN, B. P., & COLLINS, M. A. (2008). Effects of exposure to sex-stereotyped video game characters on tolerance of sexual harassment. *Journal of Experimental Social Psychology, 44*(5), 1402-1408.

DILLENBOURG, P., & JERMANN, P. (2007). Designing integrative scripts. In *Scripting computer-supported collaborative learning : Cognitive, computational and educational perspectives* (p. 275-301). Springer.

DWECK, C. S. (2000). Self-theories : their Role in Motivation. *Personality, and Development*.

EAGLY, A. H., NATER, C., MILLER, D. I., KAUFMANN, M., & SCZESNY, S. (2020). Gender stereotypes have changed : A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist, 75*(3), 301.

EAGLY, A. H., WOOD, W., & DIEKMAN, A. B. (2000). Social role theory of sex differences and similarities : A current appraisal. *The developmental social psychology of gender, 12*(174), 9781410605245-12.

EISENBERG, N., MARTIN, C. L., & FABES, R. A. (1996). Gender development and gender effects.

EKMAN, P., & FRIESEN, W. V. (1978). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.

ESPOSITO, A., AMORESE, T., CUCINIELLO, M., RIVIELLO, M. T., ESPOSITO, A. M., TRONCONE, A., TORRES, M. I., SCHLÖGL, S., & CORDASCO, G. (2021). Elder user's attitude toward assistive virtual agents : the role of voice and gender. *Journal of Ambient Intelligence and Humanized Computing, 12*(4), 4429-4436.

EYBEN, F., WÖLLMER, M., & SCHULLER, B. (2010). Opensmile : the munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM international conference on Multimedia*, 1459-1462.

FANIKO, K., ELLEMERS, N., & DERKS, B. (2021). The Queen Bee phenomenon in Academia 15 years after : Does it still exist, and if so, why? *British Journal of Social Psychology, 60*(2), 383-399.

FENG, D., JEONG, D. C., KRÄMER, N. C., MILLER, L. C., & MARSELLA, S. (2017). "Is It Just Me?" : Evaluating Attribution of Negative Feedback as a Function of Virtual Instructor's Gender and Proxemics. *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Sys*, 810-818.

FINKELSTEIN, S., YARZEBINSKI, E., VAUGHN, C., OGAN, A., & CASSELL, J. (2013). The effects of culturally congruent educational technologies on student achievement. *Artificial Intelligence in Education : 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings 16*, 493-502.

FISCHER, J.-P., & THIERRY, X. (2022). Boy's math performance, compared to girls', jumps at age 6 (in the ELFE's data at least). *British Journal of Developmental Psychology, 40*(4), 504-519.

FISCHER, M. J. (2010). A longitudinal examination of the role of stereotype threat and racial climate on college outcomes for minorities at elite institutions. *Social Psychology of Education, 13*, 19-40.

FITRIANIE, S., BRUIJNES, M., RICHARDS, D., BÖNSCH, A., & BRINKMAN, W.-P. (2020). The 19 unifying questionnaire constructs of artificial social agents : An iva community analysis. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1-8.

FLORE, P. C., MULDER, J., & WICHERTS, J. M. (2018). The influence of gender stereotype threat on mathematics test scores of Dutch high school students : A registered report. *Comprehensive Results in Social Psychology, 3*(2), 140-174.

FLORE, P. C., & WICHERTS, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of school psychology, 53*(1), 25-44.

FORD, T. E., FERGUSON, M. A., BROOKS, J. L., & HAGADONE, K. M. (2004). Coping sense of humor reduces effects of stereotype threat on women's math performance. *Personality and Social Psychology Bulletin, 30*(5), 643-653.

FORGASZ, H. J., & LEDER, G. C. (1996). Mathematics classrooms, gender and affect. *Mathematics Education Research Journal, 8*(2), 153-173.

FOX, J., & BAILENSON, J. N. (2009). Virtual virgins and vamps : The effects of exposure to female characters' sexualized appearance and gaze in an immersive virtual environment. *Sex Roles, 61*(3), 147-157.

FRANCESCHINI, G., GALLI, S., CHIESI, F., & PRIMI, C. (2014). Implicit gender–math stereotype and women's susceptibility to stereotype threat and stereotype lift. *Learning and Individual Differences, 32*, 273-277.

FRANK, M. G., EKMAN, P., & FRIESEN, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of personality and social psychology, 64*(1), 83.

FREEDMAN-DOAN, C., WIGFIELD, A., ECCLES, J. S., BLUMENFELD, P., ARBRETON, A., & HAROLD, R. D. (2000). What am I best at? Grade and gender differences in children's beliefs about ability improvement. *Journal of Applied Developmental Psychology, 21*(4), 379-402.

FRISBY, B. N., & MARTIN, M. M. (2010). Instructor–student and student–student rapport in the classroom. *Communication Education, 59*(2), 146-164.

FUMEL, S., KESKPAIK, S., SALLES, F., & VERLET, I. (2016). PISA 2015 (M. de l'éducation NATIONALE, de l'enseignement supérieur et de la recherche. DIRECTION DE L'ÉVALUATION & de la prospective et de la PERFORMANCE, Éd.) [Chaque note d'information présente les résultats les plus récents issus des exploitations d'enquêtes et d'études statistiques. Cette publication apporte l'éclairage de la direction de l'évaluation, de la prospective et de la performance (DEPP) à la compréhension et à l'analyse du système éducatif français. Elle est disponible en ligne, sous la forme de fichiers imprimables à télécharger, accompagnés de tableaux de données chiffrées et de graphiques.]. https://doi.org/10.48464/ni-16-38

GAERTNER, L., & INSKO, C. A. (2000). Intergroup discrimination in the minimal group paradigm : Categorization, reciprocation, or fear ? *Journal of personality and social psychology, 79*(1), 77.

GALDI, S., CADINU, M., & TOMASETTO, C. (2014). The roots of stereotype threat : When automatic associations disrupt girls' math performance. *Child development, 85*(1), 250-263.

GANLEY, C. M., MINGLE, L. A., RYAN, A. M., RYAN, K., VASILYEVA, M., & PERRY, M. (2013). An examination of stereotype threat effects on girls' mathematics performance. *Developmental psychology, 49*(10), 1886.

GAUTHIER, K., MORAND, A., MAZEROLLE, M., RIGALLEAU, F., EUSTACHE, F., DESGRANGES, B., HUGUET, P., & RÉGNER, I. (2020). Effet de menace du stéréotype : historique, mécanismes, et conséquences sur les performances cognitives des personnes âgées. *Revue de neuropsychologie, 12*(4), 351-357.

GILLIS, J. S., BERNIERI, F. J., & WOOTEN, E. (1995). The effects of stimulus medium and feedback on the judgment of rapport. *Organizational Behavior and Human Decision Processes, 63*(1), 33-45.

GOOD, C., ARONSON, J., & HARDER, J. A. (2008). Problems in the pipeline : Stereotype threat and women's achievement in high-level math courses. *Journal of applied developmental psychology, 29*(1), 17-28.

GOOD, C., ARONSON, J., & INZLICHT, M. (2003). Improving adolescents' standardized test performance : An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*(6), 645-662.

GOOD, C., RATTAN, A., & DWECK, C. S. (2012). Why do women opt out ? Sense of belonging and women's representation in mathematics. *Journal of personality and social psychology, 102*(4), 700.

GORDON, G., SPAULDING, S., WESTLUND, J. K., LEE, J. J., PLUMMER, L., MARTINEZ, M., DAS, M., & BREAZEAL, C. (2016). Affective personalization of a social robot tutor for children's second language skills. *Proceedings of the AAAI conference on artificial intelligence, 30*(1).

GRANITZ, N. A., KOERNIG, S. K., & HARICH, K. R. (2009). Now it's personal : Antecedents and outcomes of rapport between business faculty and their students. *Journal of Marketing Education, 31*(1), 52-65.

GRATCH, J., OKHMATOVSKAIA, A., LAMOTHE, F., MARSELLA, S., MORALES, M., van der WERF, R. J., & MORENCY, L.-P. (2006). Virtual rapport. *Intelligent Virtual Agents : 6th International Conference, IVA 2006, Marina Del Rey, CA, USA, August 21-23, 2006. Proceedings 6*, 14-27.

GRATCH, J., WANG, N., GERTEN, J., FAST, E., & DUFFY, R. (2007). Creating rapport with virtual agents. *International Workshop on Intelligent Virtual Agents*, 125-138.

GREENWALD, A. G., NOSEK, B. A., & BANAJI, M. R. (2003). Understanding and using the implicit association test : I. An improved scoring algorithm. *Journal of personality and social psychology, 85*(2), 197.

GREMLER, D. D., & GWINNER, K. P. (2000). Customer-employee rapport in service relationships. *Journal of service research, 3*(1), 82-104.

GRIVOKOSTOPOULOU, F., KOVAS, K., & PERIKOS, I. (2020). The effectiveness of embodied pedagogical agents and their impact on students learning in virtual worlds. *Applied Sciences*, *10*(5), 1739.

GUADAGNO, R. E., BLASCOVICH, J., BAILENSON, J. N., & MCCALL, C. (2007). Virtual humans and persuasion : The effects of agency and behavioral realism. *Media Psychology*, *10*(1), 1-22.

GUAY, F., VALLERAND, R. J., & BLANCHARD, C. (2000). On the assessment of situational intrinsic and extrinsic motivation : The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*(3), 175-213.

GULZ, A. (2004). Benefits of virtual characters in computer based learning environments : Claims and evidence. *International Journal of Artificial Intelligence in Education*, *14*(3-4), 313-334.

GULZ, A., HAAKE, M., & TÄRNING, B. (2007). Visual gender and its motivational and cognitive effects–a user study. *Lund University Cognitive Studies*, *137*, 928-935.

HANNA, N., & RICHARDS, D. (2018). The impact of multimodal communication on a shared mental model, trust, and commitment in human–intelligent virtual agent teams. *Multimodal Technologies and Interaction*, *2*(3), 48.

HARKINS, S. G. (2006). Mere effort as the mediator of the evaluation-performance relationship. *Journal of Personality and Social psychology*, *91*(3), 436.

HARRIGAN, J. A., & O'CONNELL, D. M. (1996). How do you look when feeling anxious? Facial displays of anxiety. *Personality and individual differences*, *21*(2), 205-212.

HART, S. G., & STAVELAND, L. E. (1988). Development of NASA-TLX (Task Load Index) : Results of empirical and theoretical research. In *Advances in Psychology* (p. 139-183, T. 52). Elsevier.

HARTLEY, P. (2002). *Interpersonal communication*. Routledge.

HAYES, A. L., ULINSKI, A. C., & HODGES, L. F. (2010). That avatar is looking at me ! Social inhibition in virtual worlds. *International Conference on Intelligent Virtual Agents*, 454-467.

HEILMAN, M. E., & OKIMOTO, T. G. (2007). Why are women penalized for success at male tasks ? : the implied communality deficit. *Journal of applied psychology*, *92*(1), 81.

HEYSELAAR, E., CARUANA, N., SHIN, M., SCHILBACH, L., & CROSS, E. S. (2023). Do we really interact with artificial agents as if they are human ? *Frontiers in Virtual Reality*, *4*, 1201385.

HIRSH-PASEK, K., ZOSH, J. M., GOLINKOFF, R. M., GRAY, J. H., ROBB, M. B., & KAUFMAN, J. (2015). Putting education in "educational" apps : Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3-34.

HORTON, D., & RICHARD WOHL, R. (1956). Mass communication and para-social interaction : Observations on intimacy at a distance. *psychiatry*, *19*(3), 215-229.

HOYT, C. L., & SIMON, S. (2011). Female leaders : Injurious or inspiring role models for women ? *Psychology of Women Quarterly*, *35*(1), 143-157.

HUANG, L. (2013). *Parasocial consensus sampling : Modeling human nonverbal behaviors from multiple perspectives* [thèse de doct., University of Southern California].

HUANG, L., MORENCY, L.-P., & GRATCH, J. (2011). Virtual Rapport 2.0. *Intelligent Virtual Agents : 10th International Conference, IVA 2011, Reykjavik, Iceland, September 15-17, 2011. Proceedings 11*, 68-79.

HUGUET, P., & REGNER, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of educational psychology*, *99*(3), 545.

HUGUET, P., & RÉGNER, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of experimental social psychology*, *45*(4), 1024-1027.

HUSTON, A. C. (1985). The development of sex typing : Themes from recent research. *Developmental Review*, *5*(1), 1-17.

HYDE, J. S., & LINN, M. C. (2006). Gender similarities in mathematics and science. *Science, 314*(5799), 599-600.

IACOBELLI, F., & CASSELL, J. (2007). Ethnic identity and engagement in embodied conversational agents. *Intelligent Virtual Agents : 7th International Conference, IVA 2007 Paris, France, September 17-19, 2007 Proceedings 7*, 57-63.

INZLICHT, M., & BEN-ZEEV, T. (2000). A threatening intellectual environment : Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological science*, *11*(5), 365-371.

JEONG, D. C., FENG, D., KRÄMER, N. C., MILLER, L. C., & MARSELLA, S. (2017). Negative feedback in your face : Examining the effects of proxemics and gender on learning. *International Conference on Intelligent Virtual Agents*, 170-183.

JOEL, D., PERSICO, A., SALHOV, M., BERMAN, Z., OLIGSCHLÄGER, S., MEILIJSON, I., & AVERBUCH, A. (2018). Analysis of human brain structure reveals that the brain "types" typical of males are also typical of females, and vice versa. *Frontiers in human neuroscience*, *12*, 399.

JOHN-HENDERSON, N. A., RHEINSCHMIDT, M. L., & MENDOZA-DENTON, R. (2015). Cytokine responses and math performance : The role of stereotype threat and anxiety reappraisals. *Journal of Experimental Social Psychology*, *56*, 203-206.

JOHNS, M., SCHMADER, T., & MARTENS, A. (2005). Knowing is half the battle : Teaching stereotype threat as a means of improving women's math performance. *Psychological science, 16*(3), 175-179.

JOHNSON, A. M., OZOGUL, G., MORENO, R., & REISSLEIN, M. (2013). Pedagogical agent signaling of multiple visual engineering representations : The case of the young female agent. *Journal of Engineering Education*, *102*(2), 319-337.

JOHNSON, H. J., BARNARD-BRAK, L., SAXON, T. F., & JOHNSON, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education*, *80*(2), 137-149.

JUDD, C. M., & PARK, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological review, 100*(1), 109.

KAN, M. Y., SULLIVAN, O., & GERSHUNY, J. (2011). Gender convergence in domestic work : Discerning the effects of interactional and institutional barriers from large-scale data. *Sociology*, *45*(2), 234-251.

KANTHARAJU, R. B., DE FRANCO, D., PEASE, A., & PELACHAUD, C. (2018). Is two better than one ? Effects of multiple agents on user persuasion. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 255-262.

KAO, D., MUBARRAT, S. T., JOSHI, A., PANDITA, S., MOUSAS, C., LIANG, H.-N., & RATAN, R. (2024). Exploring how gender-anonymous voice avatars influence women's performance in online computing group work. *International Journal of Human-Computer Studies*, *181*, 103146.

KÄTSYRI, J., FÖRGER, K., MÄKÄRÄINEN, M., & TAKALA, T. (2015). A review of empirical evidence on different uncanny valley hypotheses : support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology*, *6*, 390.

KELLER, J. (2002). Blatant stereotype threat and women's math performance : Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, *47*(3-4), 193-198.

KELLER, J. (2007). Stereotype threat in classroom settings : The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British journal of educational psychology*, *77*(2), 323-338.

KELLER, J., & DAUENHEIMER, D. (2003). Stereotype threat in the classroom : Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, *29*(3), 371-381.

KELTNER, D. (1995). Signs of appeasement : Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of personality and social psychology*, *68*(3), 441.

KESSELS, U. (2005). Fitting into the stereotype : How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European journal of psychology of education*, *20*(3), 309-323.

KHASHE, S., LUCAS, G., BECERIK-GERBER, B., & GRATCH, J. (2017). Buildings with persona : Towards effective building-occupant communication. *Computers in Human Behavior*, *75*, 607-618.

KIEFER, A. K., & SEKAQUAPTEWA, D. (2007). Implicit stereotypes and women's math performance : How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of experimental social psychology*, *43*(5), 825-832.

KIM, S., ALISON, L., & CHRISTIANSEN, P. (2020). Observing rapport-based interpersonal techniques to gather information from victims. *Psychology, public policy, and law*, *26*(2), 166.

KIM, Y. (2007). Desirable characteristics of learning companions. *International Journal of Artificial Intelligence in Education*, *17*(4), 371-388.

KIM, Y. (2013). Digital peers to help children's text comprehension and perceptions. *Journal of Educational Technology & Society*, *16*(4), 59-70.

KIM, Y. (2016). The role of agent age and gender for middle-grade girls. *Computers in the Schools*, *33*(2), 59-70.

KIM, Y., & BAYLOR, A. L. (2016). Research-based design of pedagogical agent roles : A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education*, *26*(1), 160-169.

KIM, Y., BAYLOR, A. L., & SHEN, E. (2007). Pedagogical agents as learning companions : The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, *23*(3), 220-234.

KIM, Y., & LIM, J. H. (2013). Gendered socialization with an embodied agent : Creating a social and affable mathematics learning environment for middle-grade females. *Journal of Educational Psychology*, *105*(4), 1164.

KIM, Y., & WEI, Q. (2011). The impact of learner attributes and learner choice in an agent-based environment. *Computers & Education*, *56*(2), 505-514.

KIMBALL, M. M. (1989). A new perspective on women's math achievement. *Psychological bulletin*, *105*(2), 198.

KRÄMER, N. C., KARACORA, B., LUCAS, G., DEHGHANI, M., RÜTHER, G., & GRATCH, J. (2016). Closing the gender gap in with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education*, *99*, 1-13.

KREIJNS, K., KIRSCHNER, P. A., & JOCHEMS, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments : a review of the research. *Computers in human behavior*, *19*(3), 335-353.

KRENDL, A. C., RICHESON, J. A., KELLEY, W. M., & HEATHERTON, T. F. (2008). The negative consequences of threat : A functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science*, *19*(2), 168-175.

KULMS, P., KRÄMER, N. C., GRATCH, J., & KANG, S.-H. (2011). It's in their eyes : A study on female and male virtual humans' gaze. *International Workshop on Intelligent Virtual Agents*, 80-92.

KUNDA, Z., & OLESON, K. C. (1995). Maintaining stereotypes in the face of disconfirmation : constructing grounds for subtyping deviants. *Journal of personality and social psychology*, *68*(4), 565.

KYRIACOU, C. (2010). *Effective teaching in schools theory and practice*. Oxford University Press-Children.

LAFRANCE, M., & HECHT, M. A. (1995). Why smiles generate leniency. *Personality and Social Psychology Bulletin*, *21*(3), 207-214.

LAFRANCE, M., & VIAL, A. C. (2016). Gender and nonverbal behavior.

LAKIN, J. L., & CHARTRAND, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological science*, *14*(4), 334-339.

LAUZEN, M. M., DOZIER, D. M., & HORAN, N. (2008). Constructing gender stereotypes through social roles in prime-time television. *Journal of broadcasting & electronic media*, *52*(2), 200-214.

LEE, E.-J. (2003). Effects of "gender" of the computer on informational social influence : The moderating role of task type. *International Journal of Human-Computer Studies*, *58*(4), 347-362.

LEE, J.-E. R., & NASS, C. (2012). Distinctiveness-based stereotype threat and the moderating role of coaction contexts. *Journal of Experimental Social Psychology*, *48*(1), 192-199.

LEE, J.-E. R., NASS, C. I., & BAILENSON, J. N. (2014). Does the mask govern the mind? : Effects of arbitrary gender representation on quantitative task performance in avatar-represented virtual groups. *Cyberpsychology, Behavior, and Social Networking, 17*(4), 248-254.

LEROY, N., MAX, S., & PANSU, P. (2022). Is Emma or Liam the Top Scorer in Math? The Effects of a Counter-Stereotypical Role Model on Math Achievement. *Sex Roles, 86*(9-10), 587-603.

LESKO, A. C., & CORPUS, J. H. (2006). Discounting the difficult : How high math-identified women respond to stereotype threat. *Sex Roles, 54*, 113-125.

LESLIE, S.-J., CIMPIAN, A., MEYER, M., & FREELAND, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262-265.

LEVY, G. D., SADOVSKY, A. L., & TROSETH, G. L. (2000). Aspects of young children's perceptions of gender-typed occupations. *Sex roles, 42*, 993-1006.

LEYENS, J.-P., YZERBYT, V., & SCHADRON, G. (1994). *Stereotypes and social cognition*. Sage Publications, Inc.

LI, J., KIZILCEC, R., BAILENSON, J., & JU, W. (2016). Social robots and virtual agents as lecturers for video instruction. *Computers in Human Behavior, 55*, 1222-1230.

LIANG, T.-P., CHEN, H.-Y., & TURBAN, E. (2009). Effect of personalization on the perceived usefulness of online customer services : A dual-core theory. *Proceedings of the 11th International Conference on Electronic Commerce*, 279-288.

LINDBERG, S. M., HYDE, J., PETERSEN, J., & LINN, M. (2010). Gender similarities characterize math performance. *Psychological Bulletin, 136*, 1123-1135.

LIU, S., LIU, P., WANG, M., & ZHANG, B. (2021). Effectiveness of stereotype threat interventions : A meta-analytic review. *Journal of Applied Psychology, 106*(6), 921.

LOGEL, C., ISERMAN, E. C., DAVIES, P. G., QUINN, D. M., & SPENCER, S. J. (2009). The perils of double consciousness : The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology, 45*(2), 299-312.

LOGEL, C., WALTON, G. M., SPENCER, S. J., ISERMAN, E. C., von HIPPEL, W., & BELL, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of personality and social psychology, 96*(6), 1089.

LOIDEAIN, N. N., & ADAMS, R. (2020). From Alexa to Siri and the GDPR : the gendering of virtual personal assistants and the role of data protection impact assessments. *Computer Law & Security Review, 36*, 105366.

LOMANOWSKA, A. M., & GUITTON, M. J. (2016). Online intimacy and well-being in the digital age. *Internet interventions, 4*, 138-144.

LUCAS, G. M., RIZZO, A., GRATCH, J., SCHERER, S., STRATOU, G., BOBERG, J., & MORENCY, L.-P. (2017). Reporting mental health symptoms : breaking down barriers to care with virtual human interviewers. *Frontiers in Robotics and AI, 4*, 51.

LUGRIN, B., PELACHAUD, C., & TRAUM, D. (2022). *The Handbook on Socially Interactive Agents : 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2 : Interactivity, Platforms, Application*. ACM.

LUNARDO, R., BRESSOLLES, G., et al. (2016). The interacting effect of virtual agents' gender and dressing style on attractiveness and subsequent consumer online behavior. *Journal of Retailing and Consumer Services*, *30*(100), 59-66.

LUONG, K. T., & KNOBLOCH-WESTERWICK, S. (2017). Can the media help women be better at math ? stereotype threat, selective exposure, media effects, and women's math performance. *Human Communication Research*, *43*(2), 193-213.

MACINTOSH, G. (2009). Examining the antecedents of trust and rapport in services : Discovering new interrelationships. *Journal of Retailing and Consumer Services*, *16*(4), 298-305.

MADAIO, M., LASKO, R., OGAN, A., & CASSELL, J. (2017). Using Temporal Association Rule Mining to Predict Dyadic Rapport in Peer Tutoring. *International Educational Data Mining Society*.

MAJOR, B., SPENCER, S., SCHMADER, T., WOLFE, C., & CROCKER, J. (1998). Coping with negative stereotypes about intellectual performance : The role of psychological disengagement. *Personality and social psychology bulletin*, *24*(1), 34-50.

MAKAROVA, E., AESCHLIMANN, B., & HERZOG, W. (2019). The gender gap in fields : The impact of the gender stereotype of math and science on secondary students' career aspirations. *Frontiers in Education*, *4*, 60.

MAKRANSKY, G., WISMER, P., & MAYER, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, *35*(3), 349-358.

MARTENS, A., JOHNS, M., GREENBERG, J., & SCHIMEL, J. (2006). Combating stereotype threat : The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, *42*(2), 236-243.

MARX, D. M. (2019). Fear of the known ? The effect of peer relevance and gender on women's math performance under threat. *Social Psychology of Education*, *22*, 1197-1214.

MARX, D. M., & KO, S. J. (2012). Superstars "like" me : The effect of role model similarity on performance under threat. *European Journal of Social Psychology*, *42*(7), 807-812.

MARX, D. M., MONROE, A. H., COLE, C. E., & GILBERT, P. N. (2013). No doubt about it : When doubtful role models undermine men's and women's math performance under threat. *The Journal of social psychology*, *153*(5), 542-559.

MARX, D. M., & ROMAN, J. S. (2002). Female role models : Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183-1193.

MATSUYAMA, Y., BHARDWAJ, A., ZHAO, R., ROMEO, O., AKOJU, S., & CASSELL, J. (2016). Socially-aware animated intelligent personal assistant agent. *Proceedings of the 17th annual meeting of the special interest group on discourse and dialogue*, 224-227.

MAZEROLLE, M., REGNER, I., RIGALLEAU, F., & HUGUET, P. (2020). Highlighting and Reducing the Impact of Negative Aging Stereotypes During Older Adults' Cognitive Testing. *JoVE (Journal of Visualized Experiments)*, (155), e59922.

MCDONNELL, R., JÖRG, S., HODGINS, J. K., NEWELL, F., & O'SULLIVAN, C. (2009). Evaluating the effect of motion and body shape on the perceived sex of virtual characters. *ACM Transactions on Applied Perception (TAP)*, *5*(4), 1-14.

MCGLONE, M. S., & ARONSON, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, *27*(5), 486-493.

MCINTYRE, R. B., PAULSON, R. M., & LORD, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*(1), 83-90.

MCINTYRE, R. B., PAULSON, R. M., TAYLOR, C. A., MORIN, A. L., & LORD, C. G. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology*, *41*(3), 301-311.

MELL, J., LUCAS, G., & GRATCH, J. (2017). Prestige questions, online agents, and gender-driven differences in disclosure. *International Conference on Intelligent Virtual Agents*, 273-282.

MENDES, W. B., BLASCOVICH, J., LICKEL, B., & HUNTER, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin*, *28*(7), 939-952.

MIRAULT, J., FAVIER, M., & MASCRET, N. (2023). A modern & innovative space for educational research : the" Jacques Ginestié" experimental platform. *Recherche en éducation. Colloque international en hommage à Jacques Ginestié*.

MIYAKE, A. (1999). Toward unified theories of working memory : Emerging general consensus, unresolved theoretical issues, and future research directions. *Models of working memory : Mechanisms of active maintenance and executive control*, 442-481.

MOHTADI, M. T., HAJAMI, A., & ALLALI, H. (2014). Pedagogical agent for metacognitive scaffolding in interactive learning environments. *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, 652-656.

MORBINI, F., DEVAULT, D., SAGAE, K., GERTEN, J., NAZARIAN, A., & TRAUM, D. (2014). FLoReS : a forward looking, reward seeking, dialogue manager. *Natural Interaction with Robots, Knowbots and Smartphones : Putting Spoken Dialog Systems into Practice*, 313-325.

MORENO, K. N., PERSON, N. K., ADCOCK, A. B., ECK, R., JACKSON, G. T., & MARINEAU, J. C. (2002). Etiquette and efficacy in animated pedagogical agents : The role of stereotypes. *AAAI Symposium on Personalized Agents, Cape Cod, MA*.

MORENO, R. (2005). Multimedia learning with animated pedagogical agents.

MORENO, R., & FLOWERDAY, T. (2006). Students' choice of animated pedagogical agents in science learning : A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology*, *31*(2), 186-207.

MORENO, R., MAYER, R. E., SPIRES, H. A., & LESTER, J. C. (2001). The case for social agency in computer-based teaching : Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and instruction*, *19*(2), 177-213.

MOUSAS, C., ANASTASIOU, D., & SPANTIDI, O. (2018). The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human Behavior*, *86*, 99-108.

MÜLLER, P., HUANG, M. X., & BULLING, A. (2018). Detecting low rapport during natural interactions in small groups from non-verbal behaviour. *23rd International Conference on Intelligent User Interfaces*, 153-164.

MURPHY, E., & RODRÍGUEZ-MANZANARES, M. A. (2012). Rapport in distance education. *International Review of Research in Open and Distributed Learning*, *13*(1), 167-190.

MURPHY, M. C., STEELE, C. M., & GROSS, J. J. (2007). Signaling threat : How situational cues affect women in math, science, and engineering settings. *Psychological science*, *18*(10), 879-885.

MUSSWEILER, T. (2003). Comparison processes in social judgment : mechanisms and consequences. *Psychological review*, *110*(3), 472.

MUZZATTI, B., & AGNOLI, F. (2007). Gender and mathematics : attitudes and stereotype threat susceptibility in Italian children. *Developmental psychology*, *43*(3), 747.

NAG, P., & YALÇIN, Ö. N. (2020). Gender Stereotypes in Virtual Agents. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1-8.

NASH, A., & GROSSI, G. (2007). Picking Barbie™'s brain : Inherent sex differences in scientific ability? *Journal of Interdisciplinary Feminist Thought*, *2*(1), 5.

NASS, C., & MOON, Y. (2000). Machines and mindlessness : Social responses to computers. *Journal of Social Issues*, *56*(1), 81-103.

NGUYEN, H. T. (2007). Rapport building in language instruction : A microanalysis of the multiple resources in teacher talk. *Language and Education*, *21*(4), 284-303.

NICKELS, W. G., EVERETT, R. F., & KLEIN, R. (1983). Rapport building for salespeople : A neuro-linguistic approach. *Journal of Personal Selling & Sales Management*, *3*(2), 1-7.

NICULESCU, A., HOFS, D., VAN DIJK, B., & NIJHOLT, A. (2010). How the agent's gender influence users' evaluation of a QA sy. *2010 International Conference on User Science and Engineering (i-USEr)*, 16-20.

NICULESCU, A., VAN DER SLUIS, F., & NIJHOLT, A. (2009). Feminity, masculinity and androgyny : How humans perceive the gender of anthropomorphic agents. *Proceedings of 13th International Conference on Human-Computer Interaction, Springer Verlag, Heidelberg*, 628-632.

NIEWIADOMSKI, R., & PELACHAUD, C. (2007). Model of facial expressions management for an embodied conversational agent. *Affective Computing and Intelligent Interaction : Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007 Proceedings 2*, 12-23.

NOREM, J. K., & CANTOR, N. (1986). Defensive pessimism : Harnessing anxiety as motivation. *Journal of personality and social psychology*, *51*(6), 1208.

NOSEK, B. A., BANAJI, M. R., & GREENWALD, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of personality and social psychology*, *83*(1), 44.

NOSEK, B. A., SMYTH, F. L., SRIRAM, N., LINDNER, N. M., DEVOS, T., AYALA, A., BAR-ANAN, Y., BERGH, R., CAI, H., GONSALKORALE, K., et al. (2009). National diffe-

rences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593-10597.

NUMATA, T., SATO, H., ASA, Y., KOIKE, T., MIYATA, K., NAKAGAWA, E., SUMIYA, M., & SADATO, N. (2020). Achieving affective human–virtual agent communication by enabling virtual agents to imitate positive expressions. *Scientific reports*, *10*(1), 5977.

NUNAMAKER, J. F., DERRICK, D. C., ELKINS, A. C., BURGOON, J. K., & PATTON, M. W. (2011). Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Sys*, *28*(1), 17-48.

NUSSBAUM, A. D., & STEELE, C. M. (2007). Situational disengagement and persistence in the face of adversity. *Journal of experimental social psychology*, *43*(1), 127-134.

O'BRIEN, L. T., & CRANDALL, C. S. (2003). Stereotype threat and arousal : Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*(6), 782-789.

O'BRIEN, L. T., HITTI, A., SHAFFER, E., CAMP, A. R. V., HENRY, D., & GILBERT, P. N. (2017). Improving girls' sense of fit in science : Increasing the impact of role models. *Social Psychological and Personality Science*, *8*(3), 301-309.

OCDE. (2019). What Are the OECD Principles on AI? [https://www.oecdilibrary.org/economics/what-are-the-oecd-principles-on-ai_6ff2a1c4-en].

OCHS, M. (2020). Des Agents Virtuels pour un Monde Meilleur. *Workshop sur les Affects, Compagnons artificiels et Interactions*.

OCHS, M., DING, Y., FOURATI, N., CHOLLET, M., RAVENET, B., PECUNE, F., GLAS, N., PREPIN, K., CLAVEL, C., & PELACHAUD, C. (2015). Vers des Agents Conversationnels Animés dotés d'émotions et d'attitudes sociales. *Journal d'Interaction Personne-Système*, *3*(Special Issue" the best of IHM'2013").

OCHS, M., NIEWIADOMSKI, R., & PELACHAUD, C. (2010). How a virtual agent should smile? Morphological and dynamic characteristics of virtual agent's smiles. *Intelligent Virtual Agents : 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings 10*, 427-440.

OGOLSKY, B. G., & BOWERS, J. R. (2013). A meta-analytic review of relationship maintenance and its correlates. *Journal of Social and Personal Relationships*, *30*(3), 343-367.

OSWALD, D. L., & HARVEY, R. D. (2000). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology*, *19*, 338-356.

OZOGUL, G., JOHNSON, A. M., ATKINSON, R. K., & REISSLEIN, M. (2013). Investigating the impact of pedagogical agent gender matching and learner choice on learning outcomes and perceptions. *Computers & Education*, *67*, 36-50.

PAHLKE, E., HYDE, J. S., & ALLISON, C. M. (2014). The effects of single-sex compared with coeducational schooling on students' performance and attitudes : A meta-analysis. *Psychological bulletin*, *140*(4), 1042.

PANSU, P., RÉGNER, I., MAX, S., COLÉ, P., NEZLEK, J. B., & HUGUET, P. (2016). A burden for the boys : Evidence of stereotype threat in boys' reading performance. *Journal of Experimental Social Psychology, 65*, 26-30.

PAPANGELIS, A., ZHAO, R., & CASSELL, J. (2014). Towards a computational architecture of dyadic rapport management for virtual agents. *Intelligent Virtual Agents : 14th International Conference, IVA 2014, Boston, MA, USA, August 27-29, 2014. Proceedings 14*, 320-324.

PARK, R. E., & BURGESS, E. W. (1924). *Introduction to the Science of Sociology* (T. 1). University of Chicago Press.

PAYNE, J., SZYMKOWIAK, A., ROBERTSON, P., & JOHNSON, G. (2013). Gendering the machine : Preferred virtual assistant gender and realism in self-service. *International Workshop on Intelligent Virtual Agents*, 106-115.

PECUNE, F., MURALI, S., TSAI, V., MATSUYAMA, Y., & CASSELL, J. (2019). A model of social explanations for a conversational movie recommendation system. *Proceedings of the 7th International Conference on Human-Agent Interaction*, 135-143.

PENNINGTON, C. R., LITCHFIELD, D., MCLATCHIE, N., & HEIM, D. (2019). Stereotype threat may not impact women's inhibitory control or mathematical performance : Providing support for the null hypothesis. *European Journal of Social Psychology, 49*(4), 717-734.

PERRY, S. P., & SKITKA, L. J. (2009). Making lemonade ? Defensive coping style moderates the effect of stereotype threat on women's math test performance. *Journal of Research in Personality, 43*(5), 918-920.

PEZZULLO, L. G., WIGGINS, J. B., FRANKOSKY, M. H., MIN, W., BOYER, K. E., MOTT, B. W., WIEBE, E. N., & LESTER, J. C. (2017). "Thanks Alisha, Keep in Touch" : Gender Effects and Engagement with Virtual Learning Companions. *International Conference on Artificial Intelligence in Education*, 299-310.

PICHO, K., & STEPHENS, J. M. (2012). Culture, context and stereotype threat : A comparative analysis of young Ugandan women in coed and single-sex schools. *The Journal of Educational Research, 105*(1), 52-63.

PICKERING, M. J., & GARROD, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences, 27*(2), 169-190.

PLANT, E. A., BAYLOR, A. L., DOERR, C. E., & ROSENBERG-KIMA, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education, 53*(2), 209-215.

POTDEVIN, D. (2020). *Vers des agents conversationnels animés sociaux : Quelle influence de l'intimité virtuelle sur l'expérience utilisateur et la relation-client ?* [Thèse de doct., Université Paris-Saclay].

POTDEVIN, D., CLAVEL, C., & SABOURET, N. (2018). Virtual Intimacy, this little something between us : a study about Human perception of intimate behaviors in Embodied Conversational Agents. *Proceedings of the 18th international conference on intelligent virtual agents*, 165-172.

PRONIN, E., STEELE, C. M., & ROSS, L. (2004). Identity bifurcation in response to stereotype threat : Women and mathematics. *Journal of Experimental Social Psychology, 40*(2), 152-168.

QIAO, S., WANG, Y., & LI, J. (2017). Real-time human gesture grading based on OpenPose. *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1-6.

QUINN, D. M., & SPENCER, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of social issues*, *57*(1), 55-71.

RAFFARD, S., SALESSE, R. N., BORTOLON, C., BARDY, B. G., HENRIQUES, J., MARIN, L., STRICKER, D., & CAPDEVIELLE, D. (2018). Using mimicry of body movements by a virtual agent to increase synchronization behavior and rapport in individuals with schizophrenia. *Scientific reports*, *8*(1), 17356.

RANGEL, U., & KELLER, J. (2011). Essentialism goes social : Belief in social determinism as a component of psychological essentialism. *Journal of personality and social psychology*, *100*(6), 1056.

RATAN, R., & SAH, Y. J. (2015). Leveling up on stereotype threat : The role of avatar customization and avatar embodiment. *Computers in Human Behavior*, *50*, 367-374.

RÄTY, H., VÄNSKÄ, J., KASANEN, K., & KÄRKKÄINEN, R. (2002). Parents' explanations of their child's performance in mathematics and reading : A replication and extension of Yee and Eccles. *Sex Roles*, *46*(3), 121-128.

RAVENET, B., OCHS, M., & PELACHAUD, C. (2013). From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes. *Intelligent Virtual Agents : 13th International Conference, IVA 2013, Edinburgh, UK, August 29-31, 2013. Proceedings 13*, 263-274.

RÉGNER, I., SMEDING, A., GIMMIG, D., THINUS-BLANC, C., MONTEIL, J.-M., & HUGUET, P. (2010). Individual differences in working memory moderate stereotype-threat effects. *Psychological Science*, *21*(11), 1646-1648.

RÉGNER, I., STEELE, J. R., AMBADY, N., THINUS-BLANC, C., & HUGUET, P. (2014). Our future scientists : A review of stereotype threat in girls from early elementary school to middle school. *Revue internationale de psychologie sociale*, *27*(3), 13-51.

REIS, H. T., et al. (2018). Intimacy as an interpersonal process. In *Relationships, well-being and behaviour* (p. 113-143). Routledge.

RENIERS, R. L., CORCORAN, R., DRAKE, R., SHRYANE, N. M., & VÖLLM, B. A. (2011). The QCAE : A questionnaire of cognitive and affective empathy. *Journal of personality assessment*, *93*(1), 84-95.

RICHARDS, D., ALSHARBI, B., & ABDULRAHMAN, A. (2020). Can I help you? Preferences of young adults for the age, gender and ethnicity of a Virtual Support Person based on individual differences including personality and psychological state. *Proceedings of the Australasian Computer Science Week Multiconference*, 1-10.

RICHARDS, Z., & HEWSTONE, M. (2001). Subtyping and subgrouping : Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5*(1), 52-73.

RISTORCELLI, M., GALLEGO, E., NGUY, K., PERGANDI, J.-M., CASANOVA, R., & OCHS, M. (2023). Investigating the impact of a virtual audience's gender and attitudes on

a human speaker. *Companion Publication of the 25th International Conference on Multimodal Interaction*, 363-367.

ROSENBERG-KIMA, R., BAYLOR, A., PLANT, E., & DOERR, C. (2008). Interface agents as social models for female students : The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior*, *24*(6), 2741-2756.

ROSENBERG-KIMA, R. B., PLANT, E. A., DOERR, C. E., & BAYLOR, A. L. (2010). The influence of computer-based model's race and gender on female students' attitudes and beliefs towards engineering. *Journal of Engineering Education*, *99*(1), 35-44.

ROSENTHAL, H. E., & CRISP, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality and Social Psychology Bulletin*, *32*(4), 501-511.

ROSENTHAL, H. E., CRISP, R. J., & SUEN, M.-W. (2007). Improving performance expectancies in stereotypic domains : Task relevance and the reduction of stereotype threat. *European journal of social psychology*, *37*(3), 586-597.

RUIZ, J.-P., LASSAULT, J., SPRENGER-CHAROLLES, L., RICHARDSON, U., LYYTINEN, H., & ZIEGLER, J. C. (2017). GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture. *Approche neuropsychologique des apprentissages chez l'enfant*.

RYDELL, R. J., MCCONNELL, A. R., & BEILOCK, S. L. (2009). Multiple social identities and stereotype threat : imbalance, accessibility, and working memory. *Journal of personality and social psychology*, *96*(5), 949.

SACKETT, P. R., BORNEMAN, M. J., & CONNELLY, B. S. (2008). High stakes testing in higher education and employment : Appraising the evidence for validity and fairness. *American Psychologist*, *63*(4), 215.

SAGUY, T., REIFEN-TAGAR, M., & JOEL, D. (2021). The gender-binary cycle : The perpetual relations between a biological-essentialist view of gender, gender ideology, and gender-labelling and sorting. *Philosophical Transactions of the Royal Society B*, *376*(1822), 20200141.

SAJJADI, P., ZHAO, J., WALLGRÜN, J. O., FURMAN, T., LA FEMINA, P. C., FATEMI, A., ZIDIK, Z. E., & KLIPPEL, A. (2020). The Effect of Virtual Agent Gender and Embodiment on the Experiences and Performance of Students in Virtual Field Trips. *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 221-228.

SALLES, F., & LE CAM, M. (2016). Les performances des élèves de terminale S en mathématique (M. de l'éducation NATIONALE, de l'enseignement supérieur et de la recherche. DIRECTION DE L'ÉVALUATION & de la prospective et de la PERFORMANCE, Éd.) [Chaque note d'information présente les résultats les plus récents issus des exploitations d'enquêtes et d'études statistiques. Cette publication apporte l'éclairage de la direction de l'évaluation, de la prospective et de la performance (DEPP) à la compréhension et à l'analyse du système éducatif français. Elle est disponible en ligne, sous la forme de fichiers imprimables à télécharger, accompagnés de tableaux de données chiffrées et de graphiques.]. https://doi.org/10.48464/ni-16-35

SANSONE, D. (2019). Teacher characteristics, student beliefs, and the gender gap in STEM fields. *Educational Evaluation and Policy Analysis*, *41*(2), 127-144.

SCHACHTER, S. (1959). The psychology of affiliation : Experimental studies of the sources of gregariousness.

SCHLIMBACH, R., & KHOSRAWI-RAD, B. (2022). Towards Ethical Design Features for Pedagogical Conversational Agents.

SCHMADER, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of experimental social psychology*, *38*(2), 194-201.

SCHMADER, T., & JOHNS, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440.

SCHMADER, T., JOHNS, M., & BARQUISSAU, M. (2004). The costs of accepting gender differences : The role of stereotype endorsement in women's experience in the math domain. *Sex roles*, *50*, 835-850.

SCHMADER, T., JOHNS, M., & FORBES, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological review*, *115*(2), 336.

SCHROEDER, N. L., ADESOPE, O. O., & GILBERT, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research*, *49*(1), 1-39.

SCHWARTZ, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The journal of the learning sciences*, *4*(3), 321-354.

SCZESNY, S., NATER, C., & EAGLY, A. H. (2018a). Agency and communion : Their implications for gender stereotypes and gender identities. In *Agency and communion in social psychology* (p. 103-116). Routledge.

SCZESNY, S., NATER, C., & EAGLY, A. H. (2018b). Agency and communion : Their implications for gender stereotypes and gender identities. In *Agency and Communion in Social Psychology* (p. 103-116). Routledge.

SEBANC, A. M., PIERCE, S. L., CHEATHAM, C. L., & GUNNAR, M. R. (2003). Gendered social worlds in preschool : Dominance, peer acceptance and assertive social skills in boys' and girls' peer groups. *Social Development*, *12*(1), 91-106.

SELFHOUT, M., DENISSEN, J., BRANJE, S., & MEEUS, W. (2009). In the eye of the beholder : perceived, actual, and peer-rated similarity in personality, communication, and friendship intensity during the acquaintanceship process. *Journal of personality and social psychology*, *96*(6), 1152.

SHAFFER, E. S., MARX, D. M., & PRISLIN, R. (2013). Mind the gap : Framing of women's success and representation in STEM affects women's math performance under threat. *Sex roles*, *68*, 454-463.

SHERMAN, A. M., & ZURBRIGGEN, E. L. (2014). "Boys can be anything" : Effect of Barbie play on girls' career cognitions. *Sex roles*, *70*, 195-208.

SHIBAN, Y., SCHELHORN, I., JOBST, V., HÖRNLEIN, A., PUPPE, F., PAULI, P., & MÜHLBERGER, A. (2015). The appearance effect : Influences of virtual agent features on performance and motivation. *Computers in Human Behavior*, *49*, 5-11.

SHIBLEY HYDE, J., & KLING, K. C. (2001). Women, motivation, and achievement. *Psychology of women quarterly*, *25*(4), 364-378.

SHIH, M., PITTINSKY, T. L., & AMBADY, N. (1999). Stereotype susceptibility : Identity salience and shifts in quantitative performance. *Psychological science*, *10*(1), 80-83.

SHRUM, W., CHEEK JR, N. H., & MACD, S. (1988). Friendship in school : Gender and racial homophily. *Sociology of Education*, 227-239.

SILVERVARG, A., HAAKE, M., & GULZ, A. (2013). Educational potentials in visually androgynous pedagogical agents. *International Conference on Artificial Intelligence in Education*, 599-602.

SINHA, T. (2016). Similarity in observable behaviors : A synthesis of studies with implications for socially-aware educational technology design. *arXiv preprint arXiv :1608.02977*.

SINHA, T., & CASSELL, J. (2015). We click, we align, we learn : Impact of influence and convergence processes on student learning and rapport building. *Proceedings of the 1st workshop on modeling INTERPERsonal SynchrONy And influence*, 13-20.

SMEDING, A., DUMAS, F., LOOSE, F., & RÉGNER, I. (2013). Order of administration of math and verbal tests : An ecological intervention to reduce stereotype threat on girls' math performance. *Journal of Educational Psychology*, *105*(3), 850.

SMITH, J. L., & WHITE, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance : It's not just a woman's issue. *Sex roles*, *47*, 179-191.

SOLIMAN, M., & GUETL, C. (2013). Implementing Intelligent Pedagogical Agents in virtual worlds : Tutoring natural science experiments in OpenWonderland. *2013 IEEE Global Engineering Education Conference (EDUCON)*, 782-789.

SPENCER, S. J., LOGEL, C., & DAVIES, P. G. (2016). Stereotype threat. *Annual review of psychology*, *67*, 415-437.

SPENCER, S. J., STEELE, C. M., & QUINN, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4-28.

SPILIOTO-POULOS, D., MAKRI, E., VASSILAKIS, C., & MARGARIS, D. (2020). Multimodal Interaction : Correlates of Learners' Metacognitive Skill Training Negotiation Experience. *Information*, *11*(8), 381.

STEELE, C. M. (1997). A threat in the air : How stereotypes shape intellectual identity and performance. *American psychologist*, *52*(6), 613.

STEELE, C. M. (2011). *Whistling Vivaldi : How stereotypes affect us and what we can do*. WW Norton & Company.

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.

STEELE, C. M., & ARONSON, J. A. (2004). Stereotype threat does not live by Steele and Aronson (1995) alone.

STEELE, C. M., & DAVIES, P. G. (2003). Stereotype threat and employment testing : A commentary. *Human Performance*, *16*(3), 311-326.

STEELE, J. (2003). Children's gender stereotypes about math : The role of stereotype stratification 1. *Journal of applied social psychology*, *33*(12), 2587-2606.

STOKES, J., RIGER, S., & SULLIVAN, M. (1995). Measuring perceptions of the working environment for women in corporate settings. *Psychology of Women Quarterly*, *19*(4), 533-549.

STOUT, J. G., DASGUPTA, N., HUNSINGER, M., & MCMANUS, M. A. (2011). STEMing the tide : using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of personality and social psychology*, *100*(2), 255.

STRENGERS, Y., & KENNEDY, J. (2021). *The smart wife : Why Siri, Alexa, and other smart home devices need a feminist reboot*. Mit Press.

STRICKER, L. J., & BEJAR, I. I. (2004). Test difficulty and stereotype threat on the GRE general test 1. *Journal of Applied Social Psychology*, *34*(3), 563-597.

SUMMERS, L. H. (2005). Remarks at NBER conference on diversifying the science & engineering workforce. *The Office of the President, Harvard University*.

ŚWIDRAK, J., POCHWATKO, G., & INSABATO, A. (2021). Does an agent's touch always matter? Study on virtual Midas touch, masculinity, social status, and compliance in Polish men. *Journal on Multimodal User Interfaces*, *15*(2), 163-174.

TAILLANDIER-SCHMITT, A., ESNARD, C., & MOKOUNKOLO, R. (2012). Self-affirmation in occupational training : Effects on the math performance of French women nurses under stereotype threat. *Sex Roles*, *67*, 43-57.

TAYLOR, C. A., LORD, C. G., MCINTYRE, R. B., & PAULSON, R. M. (2011). The Hillary Clinton effect : When the same role model inspires or fails to inspire improved performance under stereotype threat. *Group Processes & Intergroup Relations*, *14*(4), 447-459.

ter STAL, S., TABAK, M., op den AKKER, H., BEINEMA, T., & HERMENS, H. (2020). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human–Computer Interaction*, *36*(9), 881-892.

TESSER, A. (1986). Some effects of self-evaluation maintenance on cognition and action.

TETTEGAH, S., & ANDERSON, C. J. (2007). Pre-service teachers' empathy and cognitions : Statistical analysis of text data by graphical models. *Contemporary Educational Psychology*, *32*(1), 48-82.

TICKLE-DEGNEN, L., & ROSENTHAL, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, *1*(4), 285-293.

TOMASETTO, C., ALPARONE, F. R., & CADINU, M. (2011). Girls' math performance under stereotype threat : the moderating role of mothers' gender stereotypes. *Developmental psychology*, *47*(4), 943.

TSUI, A. B. (1996). Reticence and Anxiety in Second Language Learning. *Voices From the Language Classroom*.

van der LUBBE, L. M., & BOSSE, T. (2017). Studying gender bias and social backlash via simulated negotiations with virtual agents. *International Conference on Intelligent Virtual Agents*, 455-458.

VAN HONK, J., TERBURG, D., & BOS, P. A. (2011). Further notes on testosterone as a social hormone.

VAN LOO, K. J., BOUCHER, K. L., RYDELL, R. J., & RYDELL, M. T. (2013). Competition in stereotyped domains : Competition, social comparison, and stereotype threat. *European Journal of Social Psychology*, *43*(7), 648-660.

VANDELLO, J. A., BOSSON, J. K., COHEN, D., BURNAFORD, R. M., & WEAVER, J. R. (2008). Precarious manhood. *Journal of personality and social psychology*, *95*(6), 1325.

VICENTE, L., & MATUTE, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, *13*(1), 15737.

VILARO, M. J., WILSON-HOWARD, D. S., NEIL, J. M., TAVASSOLI, F., ZALAKE, M. S., LOK, B. C., MODAVE, F. P., GEORGE, T. J., ODEDINA, F. T., CAREK, P. J., et al. (2021). A Subjective Culture Approach to Cancer Prevention : Rural Black and White Adults' Perceptions of Using Virtual Health Assistants to Promote Colorectal Cancer Screening. *Health Communication*, 1-12.

von HIPPEL, C., WALSH, A. M., & ZOUROUDIS, A. (2011). Identity separation in response to stereotype threat. *Social Psychological and Personality Science*, *2*(3), 317-324.

WALTON, G. M., & COHEN, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*(5), 456-467.

WANG, I., & RUIZ, J. (2021). Examining the use of nonverbal communication in virtual agents. *International Journal of Human–Computer Interaction*, *37*(17), 1648-1673.

WANG, N., & GRATCH, J. (2009). Rapport and facial expression. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1-6.

WEBSTER, J., & WATSON, R. T. (2002). Analyzing the past to prepare for the future : Writing a literature review. *MIS Quarterly*, xiii-xxiii.

WEGNER, D. M. (1994). Ironic processes of mental control. *Psychological review*, *101*(1), 34.

WELLER, M. (2018). Twenty years of Edtech. *Educause Review Online*, *53*(4), 34-48.

WEST, M., KRAUT, R., & EI CHEW, H. (2019). I'd blush if I could : Closing gender divides in digital skills through education.

WIGFIELD, A., & ECCLES, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary educational psychology*, *25*(1), 68-81.

WIRZBERGER, M., SCHMIDT, R., GEORGI, M., HARDT, W., BRUNNETT, G., & REY, G. D. (2019). Effects of system response delays on elderly humans' cognitive performance in a virtual training scenario. *Scientific Reports*, *9*(1), 1-12.

WOUT, D., DANSO, H., JACKSON, J., & SPENCER, S. (2008). The many faces of stereotype threat : Group-and self-threat. *Journal of experimental social psychology*, *44*(3), 792-799.

WRAGA, M., HELT, M., JACOBS, E., & SULLIVAN, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social cognitive and affective neuroscience*, *2*(1), 12-19.

WYLIE, L. (2024). Education App Revenue and Usage Statistics [https://www.businessofapps.com/data/education-app-market/].

YEE, D. K., & ECCLES, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles, 19*(5), 317-333.

YEE, N., & BAILENSON, J. (2007). The Proteus effect : The effect of transformed self-representation on behavior. *Human communication research, 33*(3), 271-290.

ZAJONC, R. B. (1965). Social Facilitation : A solution is suggested for an old unresolved social psychological problem. *Science, 149*(3681), 269-274.

ZANBAKA, C., GOOLKASIAN, P., & HODGES, L. (2006). Can a virtual cat persuade you ? The role of gender and realism in speaker persuasiveness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1153-1162.

ZANDER, L., HÖHNE, E., HARMS, S., & PFOST, M. (2020). When grades are high but self-efficacy is low : Unpacking the confidence gap between girls and boys in mathematics. *Frontiers in psychology, 11*, 552355.

ZHAO, R., PAPANGELIS, A., & CASSELL, J. (2014). Towards a dyadic computational model of rapport management for human-virtual agent interaction. *Intelligent Virtual Agents : 14th International Conference, IVA 2014, Boston, MA, USA, August 27-29, 2014. Proceedings 14*, 514-527.

ZHAO, Z., MADAIO, M., PECUNE, F., MATSUYAMA, Y., & CASSELL, J. (2018). Socially-conditioned task reasoning for a virtual tutoring agent. *Proceedings of the 17th international conference on autonomous agents and MultiAgent systems*, 2265-2267.

ZHOU, X. (2021). Toward the positive consequences of teacher-student rapport for students' academic engagement in the practical instruction classrooms. *Frontiers in psychology, 12*, 759785.

ZIBREK, K., NIAY, B., OLIVIER, A.-H., HOYET, L., PETTRÉ, J., & MCDONNELL, R. (2020). The effect of gender and attractiveness of motion on proximity in virtual reality. *ACM Transactions on Applied Perception (TAP), 17*(4), 1-15.

ANNEXES

A. Liste des acronymes

- ACA : Agent Conversationnel Animé.
- Ampiric : Aix-Marseille – Pôle d'Innovation, de Recherche, d'enseIgnement pour l'éduCation.
- CRF : *Conditional Random Fields*.
- HTML : *HyperText Markup Language*.
- LSTM : *Long Short-Term Memory*.
- OCDE : Organisation de Coopération et de Développement Économiques.
- PHP : *PHP Hypertext Preprocessor*.
- PISA : *Programme for International Student Assessment*.
- REP+ : Réseau d'Education Prioritaire renforcé.
- SSML : *Speech Synthesis Markup Language*.
- STIM : Science, Technologie, Ingénierie et Mathématiques.
- TAI : Test des Associations Implicites.
- TIMSS : *Trends in International Mathematics and Science Study*.
- XML : *eXtensible Markup Language*.

B. Autorisation du comité d'éthique

Comité d'éthique de l'université d'Aix-Marseille

Objet : Avis du Comité d'éthique.
N/Réf dossier : 2021-09-30-04
Dossier suivi par :DRV-Audrey Janssens
Pièce(s) jointe(s) : 1 document

Marseille, le mardi 5 octobre 2021

Le projet de recherche présenté par l'investigateur principal, Ochs Magalie, Maître de Conférence au Laboratoire LIS de l'Université d'Aix-Marseille et l'investigatrice secondaire, Régner Isabelle, Professeur des Universités, intitulé « Améliorer les performances des enfants en mathématiques: un compagnon virtuel d'apprentissagepour lutter contre les effets de menace du stéréotype» a été soumis pour avis au Comité d'éthique en sa séance du jeudi 30 septembre 2021.

Après audition des rapporteurs, le comité a jugé que le projet ne pose pas de problème éthique ou règlementaire

Le Comité d'éthique de l'Université d'Aix-Marseille émet donc un avis favorable.

Le Président du Comité d'éthique
Pierre-Jean Weiller

Université d'Aix-Marseille - Jardin du Pharo - 58 boulevard Charles Livon - 13284 Marseille cedex 07 - France
Tél. : +33 (0)4 91 39 65 00 - Fax : +33 (0)4 91 31 31 36 - www.univ-amu.fr

C. Formulaire d'autorisation parentale

FORMULAIRE RECUEIL DU CONSENTEMENT DESTINEE AUX REPRESENTANTS D'UN ENFANT MINEUR

« Améliorer les performances des enfants en mathématiques grâce à un compagnon virtuel d'apprentissage. »

Madame, Monsieur,

Vous avez accepté d'envisager que votre enfant **(nom prénom.....)** participe à la présente étude.

POURQUOI FAIRE CETTE EXPERIENCE ?

L'étude à laquelle votre enfant va participer est destinée à mieux comprendre comment l'on peut améliorer l'apprentissage des mathématiques des enfants de 6e et 5e, sur un ordinateur, grâce à un compagnon virtuel d'apprentissage.

COMMENT SE DEROULE L'EXPERIENCE ?

Si vous acceptez cette étude, votre enfant passera deux exercices de mathématiques en deux séances séparées, sur ordinateur.

QUELS SONT LES BENEFICES ATTENDUS ?

Cette recherche est sans bénéfice individuel direct, c'est-à-dire qu'elle n'apportera pas d'information sur les capacités de votre enfant. Elle n'a pas non plus de valeur de test médical sur le développement de votre enfant.

QUELLES SONT LES CONTRAINTES ET DESAGREMENTS ?

Les exercices de maths prévus sont ceux habituellement réalisés dans le cadre du programme. Ils seront effectués sur ordinateur, dans une salle informatique du collège. L'étude est conçue avec plusieurs équipes du CNRS, les équipes enseignantes et la direction des établissements scolaires, pour être en adéquation avec le programme de mathématiques, et ne perturbera pas le cours normal des apprentissages.

QUELS SONT VOS DROITS EN TANT QUE PARTICIPANT(E) A CETTE RECHERCHE ?

La participation de votre enfant est basée sur le volontariat, il peut refuser de participer à cette étude ou vous pouvez refuser qu'il y participe sans fournir aucune forme de justification. Si c'est le cas, votre enfant réalisera les exercices de mathématiques sur feuille, comme d'habitude. Les modalités de ce protocole ont été soumises à avis et autorisation auprès du comité d'éthique d'Aix-Marseille Université, lequel a pour mission de vérifier les conditions requises pour la protection de votre enfant et le respect de ses droits : Les données recueillies seront stockées sur fichiers informatiques et seront anonymes (le nom de votre enfant ne sera pas enregistré dans le fichier, il sera impossible pour les investigateurs de cette recherche de connaître votre enfant).

Ce document vous appartient et nous vous invitons à en discuter avec votre enfant et vos proches.

SIGNATURE DU REPRESENTANT LEGAL DE L'ENFANT :

NOM :..... **Prénom :**..... **Date de naissance :**.....

Adresse : .....

Téléphone(s) : .....

Signature, précédée de la mention « lu et approuvé » :

FO-DRV-1003

D. Dialogues de l'agent avec rapport, avec les balises SSML

Présentation de l'agent

Bonjour!
Et <emphasis>bienvenue</emphasis> dans ton cours de <prosody pitch= "medium"> maths</prosody> sur<break time="10ms"/>ordinateur.
Je suis <emphasis>là</emphasis> pour t'aider à revoir <prosody pitch="high">ta </prosody> <emphasis>leçon</emphasis> <break time="50ms"/>sur

- **[Pour les 6e]** le parallélisme.
- **[Pour les 5e]** les <emphasis>priorités</emphasis> opératoires.

C'est quoi ton <prosody pitch="high">nom?</prosody>
[L'élève tape son prénom]
Sympa<break time="10ms"/> ton prénom.

- **[Agent féminin]** Moi c'est <prosody rate="high" pitch="+2st">Nora</prosody>.
- **[Agent masculin]** Moi c'est Dylan.

Je suis au <emphasis>lycéé</emphasis>, et comme toi, j'ai appris les maths au collège.

Ça va? Tu t'en sors bien en <emphasis>maths</emphasis>?
[L'élève a 4 choix de réponse.]

- **[Choix 1 : "Oui!"]** Ah oui? <emphasis>super!</emphasis> Moi aussi <break time="20ms"/>je suis comme <emphasis>toi,</emphasis>
- **[Choix 2 : "Des fois oui, des fois non."]** <prosody pitch="medium" rate="high">Oui je te <emphasis>comprends</emphasis></prosody>, parce que moi aussi<break time="20ms"/> ça dépendait des fois. Des fois,<break time="20ms">je me <emphasis>plantais</emphasis> <emphasis>totalement</emphasis> <break time= "50ms"/> et des fois ça allait. Mais maintenant,
- **[Choix 3 : "Bof."]** <prosody pitch="medium" rate="high"> Oui je te <emphasis>comprends</emphasis></pro moi au<emphasis>ssi</emphasis> <break time="20ms"/>au début j'étais comme toi, c'était bof. Mais maintenant,
- **[Choix 4 : "Non!"]** <prosody pitch="high">T'inquiète</prosody>, moi aussi <break time="50ms"/>j'avais du mal avec les <emphasis>maths</emphasis> et je trouvais ça compliqué. j'avais des notes <emphasis>complètement</emphasis> <break time="20ms"/> catastrophique. Mais maintenant,

j'ai des bonnes <emphasis>notes</emphasis> en maths, parce que j'ai passé beaucoup de temps à <emphasis>réviser</emphasis> et à faire mes devoir.
Du coup j'ai mieux compris<break time="20ms"/>les leçons et les exercices.
Y'a pas de secret en fait, on me l'avait dit, <emphasis>mais c'est vraiment vrai </emphasis>, faut réviser et écouter en cours, et après ça marche.
Allez, <break time="100ms"/>on va revoir le cours ensemble, <break time="50ms"/>tu <emphasis>verras ça va être bien</emphasis>.
Faudra que tu passes un test de maths à la fin.
Je suis sûr que tu vas bien réussir.

Leçon de 6e

Je vais te parler des droites sécantes, des droites parallèles, et des droites perpendiculaires.
On va <emphasis>comencer</emphasis> par les <emphasis>droites sécantes </emphasis>.
Deux droites sont <emphasis>sécantes</emphasis> quand elles se croisent.
Tu peux voir dans cet <emphasis>exemple</emphasis>, que les deux droites <emphasis>sécantes</emphasis>, se croisent au point E.
Ce point E, on l'appelle le point d'intersection, parce qu'il y a <break time="50ms"/> l'intersection <break time="100ms"/> des 2 droites.
Les mots à retenir sont <emphasis>droites sécantes</emphasis> <break time="50ms"/> pour dire que 2 droites se <emphasis>croisent</emphasis>,
et <emphasis>point d'intersection</emphasis> <break time="50ms"/>pour dire où elles se croisent.
Je te donne un autre exemple.
Ici tu as 2 droites. <prosody rate="high">On dirait qu'elles se croisent pas, mais si on continuait à tracer les droites, elles finiraient par se croiser.

Les droites sécantes, c'est un peu le <emphasis>contraire</emphasis> des droites <emphasis>parallèles</emphasis>, parce que 2 droites parallèles se croisent jamais.
Regarde, ici la <emphasis>droite</emphasis> D <break time="100ms"/>est <emphasis>parallèle</emphasis> à la droite D'.
Et dans cet exemple aussi, la droite D <break time="100ms"/>est <emphasis>parallèle</emphasis> à la droite D'.
On note comme ça, pour dire que les droites sont <emphasis>parallèles</emphasis> entre elles.
On a aussi les droites <break time="50ms"/>perpendiculaires.
C'est quand elles sont <emphasis>sécantes</emphasis> en formant un angle droit.

Regarde,
dans cet <emphasis>exemple</emphasis>,
la droite <break time="100ms"/>D<break time="200ms"/> est <emphasis> perpendiculaire</emphasis> à la droite <break time="100ms"/>F.
Tu vois ?
Elles se croisent <break time="50ms"/>et ça fait un angle droit.
On note comme ça, pour dire que les droites sont <emphasis>perpendiculaires </emphasis> <break time="100ms"/> entre elles.
voila, on a revu 3 définitions.
Ce que tu dois retenir,
c'est que des droites <emphasis>sécantes</emphasis> se <emphasis>croisent </emphasis>,
des droites <emphasis>parallèles</emphasis> se croisent <emphasis>jamais </emphasis>,
et des droites <emphasis>perpendiculaires</emphasis> se croisent<break time="200ms"/> en formant un angle droit.
On va continuer avec <emphasis>3</emphasis> propriétés.

Ça va, tu arrives à suivre ?
[L'élève a 2 choix.]

— **[Choix 1 : "Oui."]** J'en étais <emphasis>sûr</emphasis>, ça fait plaisir. Alors on va revoir ensemble les propriétés.

— **[Choix 2 : "Pas vraiment."]** T'inquiète, on va revoir tout ça avec les propriétés.

La propriété 1<break time="100ms"/> sert à montrer que 2 droites sont <emphasis>parallèles</emphasis>, et donc <break time="200ms"/>qu'elles se croisent jamais.
Si 2 droites sont <emphasis>perpendiculaires</emphasis> à une même troisième droite,<break time="100ms"/> alors ces deux droites sont <emphasis>parallèles</emphasis> entre elles.
Ici, on sait que la droite <break time="100ms"/>D3, est <emphasis>perpendiculaire </emphasis> à la droite <break time="100ms"/>D1, et aussi à la droite <break time="200ms"/> D2.
Donc, D1, et, D2, sont <emphasis>parallèles</emphasis> entre elles.

Il y a une autre façon de montrer que 2 droites sont <emphasis>parallèles</emphasis>, c'est la propriété 2.
Si 2 droites sont <emphasis>parallèles</emphasis> à une <emphasis>même </emphasis> troisième droite, <break time="200ms"/> alors elles sont <emphasis>parallèles </emphasis> entre elles.
Regarde,
ici,
on sait que la droite <break time="100ms"/>D1 est <emphasis>parallèle</emphasis> à la droite <break time="100ms"/>D2, et on sait aussi, que la droite D<emphasis>3</emphasis> est <emphasis>paralèle</emphasis> à la droite D2.
Du coup, on peut conclure grâce à la propriété 2 que D1 et D3 sont <emphasis>parallèles</emphasis>, car ce sont 2 droites qui sont <emphasis>parallèles</emphasis> à une même <emphasis>troisième</emphasis> droite.

La dernière <emphasis>propriété</emphasis>, sert à montrer que 2 droites sont perpendiculaires.
Tu sais, c'est quand 2 droites se croisent en formant un angle droit.
La propriété 3, est que si 2 droites sont parallèles, et si une troisième droite est <emphasis>perpendiculaire</emphasis> à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
Tu vois, ici, on sait que les droites D1 et D2 sont parallèles, et on sait que D3 est perpendiculaire à D2.
Du coup on peut dire que d3 est perpendiculaire à D1.

J'ai fini le cours.
Je te donne quelques conseils pour te facilité la tâche.
Déjà, lis bien l'énoncé pour voir de quelle droite on parle.
Et aussi pour voir toute les informations qu'on a.
Et enfin, quand tu vois une figure, regarde à quelle figure-clé elle ressemble parmi les 3 qu'on a vu.
Je te les remets ici.
Utilise la <emphasis>propriété</emphasis> qui correspond à une de ces figures <emphasis>clés</emphasis>, pour répondre aux questions.

Leçon de 5e

Je vais te parler du <emphasis>vocabulaire</emphasis> des opérations,

et ensuite des <emphasis>méthodes</emphasis> pour savoir <emphasis>calculer </emphasis> une expression.
On va <emphasis>comencer</emphasis> par le vocabulaire.
Tu as <emphasis>4</emphasis> opérations.
L'addition, la soustraction, la division, et la multiplication.
Quand tu as une addition, le résultat s'appelle la somme.
Et quand tu as une soustraction, le résultat s'appelle la différence.
Pour la division, le résultat c'est le quotient, et pour la multiplication, le résultat c'est le produit.

Pour savoir comment <emphasis>calculer</emphasis> une expression, tu dois savoir par quelle <emphasis>opération</emphasis> il faut commencer.
Je vais d'abord te parler des <emphasis>expressions</emphasis> <emphasis>sans </emphasis> parenthèse. Tu dois calculer de la <emphasis>gauche</emphasis> vers la droite.
Mais, il y a des opération qu'il faut faire en premier!
On calcule d'abord la multiplication ou la division, puis l'addition ou la soustraction.
Par exemple,
on part de la gauche, et on calcule d'abord la multiplication.
On a donc, A égal à 3+18.
On a plus qu'à <emphasis>additionner</emphasis> 3 et 18 pour avoir le résultat.
Donc A, est égal à 21.

Quand il y a des parenthèses,
tu dois d'abord calculer ce qu'il y a dans les parenthèses les plus intérieures,
tout en respectant les ordres des opérations.
Tu sais,
d'abord la multiplication ou la division, ensuite l'addition ou la soustraction.
Par exemple,
dans cette expression,
On calcule <emphasis>d'abord</emphasis> ce qu'il y a dans la
<emphasis>parenthèse</emphasis> la plus intérieure.
On a donc entre parenthèse 3+7, multiplié par 5.
Il nous reste une parenthèse. On calcule donc ce qu'il y a dedans.
Il nous reste plus qu'à <emphasis>calculer<emphasis> cette multiplication.
Et on a, B, égal a 50.

Ça va, tu arrives à suivre?
[L'élève a 2 choix.]

- **[Choix 1 : "Oui."]** J'en étais <emphasis>sûr</emphasis>, ça fait plaisir. Je vais te montrer des exemples, et après on aura fini avec la leçon.
- **[Choix 2 : "Pas vraiment."]** T'inquiète, on va revoir tout ça avec des exemples.

Au début, je t'ai parlé du <emphasis>vocabulaire</emphasis> des résultats des <emphasis>opérations</emphasis>, avec la somme pour l'addition, la différence pour la soustraction, le quotient pour la division, et le produit pour la multiplication.
Maintenant, il faut savoir si une expression est une somme, une différence, un quotient, ou un produit.
Et pour ça, il faut que tu vois quelle est la dernière opération qu'il faut calculer. Je vais te

donner 2 exemples.
Ici, on a A, égal à 3+4 multiplié par 5.
Quand tu calcules ça, tu fais d'abord la multiplication 4 fois 5, parce qu'on commence par la multiplication.
Ça te fait, A égale à 3 plus 20.
Pour finir cette <emphasis>expression</emphasis>, tu fais une addition.
Donc, A, est une somme, parce que tu as <emphasis>fini</emphasis> par une addition.

Maintenant, je vais te montrer un exemple avec des parenthèses.
Tu sais ? Quand il y a des parenthèses, il faut commencer par là.
Ici, on a, B, égal à <emphasis>20<emphasis> divisé par 7 moins 2, entre parenthèses.
Comme il faut commencer par ce qu'il y a dans les parenthèses, tu fais <emphasis>d'abord</emphasis> 7 moins 2.
Ça te fait, B, égal à 20 divisé par 5.
Tu finis ton calcul en faisant une <emphasis>division</emphasis>, du coup, B, est un quotient.

Voilà, j'ai fini le cours !
Je te donne quelques conseils pour te faciliter la tâche.
Quand tu veux <emphasis>calculer</emphasis> une expression, regarde ce qu'il faut calculer en premier.
N'oublie pas, tu calcules <emphasis>d'abord</emphasis> ce qu'il y a dans les parenthèses, et tu appliques les règles des <emphasis>priorités</emphasis> opératoires.
D'abord la multiplication ou la division, ensuite, l'addition ou la soustraction.
Et il faut calculer de gauche à droite.
Par contre, quand on ne te demande <emphasis>pas</emphasis> le résultat de l'expression, et qu'on te demande si c'est une somme, ou une différence, ou un quotient, ou un produit, il faut que tu arrives à voir la <emphasis>dernière</emphasis> opération que tu dois faire.

Fin du discours

Voilà, je vais te laisser faire le test de maths.
[Le discours change selon ce que l'élève avait répondu à la question "Ça va ? Tu t'en sors bien en maths ?" précédemment.]

- **[Choix 1 : "Oui !"]** Je suis <emphasis>sûr</emphasis> que tu vas y arriver !
- **[Choix 2 : "Des fois oui, des fois non."]** <emphasis>Même</emphasis> si tu m'as dit que des fois, ça allait pas trop en maths, je suis sûr que tu vas y arriver !
- **[Choix 3 : "Bof."]** <emphasis>Même</emphasis> si tu m'as dit que tu t'en sortais pas <emphasis>trop</emphasis> en maths, je suis sûr que tu vas y arriver !
- **[Choix 4 : "Non !"]** Même si tu m'as dit que tu t'en sortais pas en maths, je suis sûr que tu vas y arriver !

E. Dialogues de l'agent sans rapport, avec les balises SSML

Présentation de l'agent

Bonjour.
Et <emphasis>bienvenue</emphasis> dans ton cours de <prosody pitch= "medium"> maths</prosody> sur<break time="10ms"/>ordinateur.
Je suis <emphasis>là</emphasis> pour t'aider à revoir <prosody pitch="high">ta</prosody> <emphasis>leçon</emphasis> <break time="50ms"/>sur

- [**Pour les 6e**] le parallélisme.
- [**Pour les 5e**] les <emphasis>priorités</emphasis> opératoires.
- [**Agent féminin**] Je m'appelle Nora.
- [**Agent masculin**] Je m'appelle Dylan.

Je suis au <emphasis>lycéé</emphasis>, et j'ai appris les maths au collège.

- [**Agent avec explication.**] Comme j'ai passé beaucoup de temps à <emphasis>réviser</emphasis> et à faire mes devoir, j'ai mieux compris<break time="20ms"/> les leçons et les exercices.
  Maintenant j'ai des bonnes notes en maths.
  Y'a pas de secret en fait, on me l'avait dit, <emphasis>mais c'est vraiment vrai</emphasis>, faut réviser et écouter en cours, et après ça marche.

on va revoir le cour ensemble, et après tu passeras un test de maths.

Leçon de 6e

On va <emphasis>commencer</emphasis> par les <emphasis>droites sécantes </emphasis>.
Deux droites sont <emphasis>sécantes</emphasis> quand elles se croisent.
Dans cet exemple,
les 2 droites <emphasis>sécantes</emphasis> se croisent au point E.
Ce point E, on l'appelle le point d'intersection, parce qu'il y a <break time="50ms"/> l'intersection <break time="100ms"/> des 2 droites.
Les mots à retenir sont <emphasis>droites sécantes</emphasis> <break time="50ms"/> pour dire que 2 droites se <emphasis>croisent</emphasis>,
et <emphasis>point d'intersection</emphasis> <break time="50ms"/>pour dire où elles se croisent.
Voici un autre exemple.
Ici tu as 2 droites.
<prosody rate="high">on dirait qu'elles se croisent pas, mais si on augmentait la taille de l'image, elles finiraient par se croiser</prosody>.

Les droites sécantes, c'est un peu le <emphasis>contraire</emphasis> des droites <emphasis>parallèles</emphasis>, parce que 2 droites parallèles se croisent jamais.
Ici la <emphasis>droite</emphasis> D <break time="100ms"/>est <emphasis>parallèle </emphasis> à la droite D'.
Et dans cet exemple aussi, la droite D <break time="100ms"/>est <emphasis>parallèle </em-

phasis> à la droite D'.
On note come ça,
pour dire que les droites sont <emphasis>parallèles</emphasis> entre elles.
On a aussi les droites <break time="50ms"/>perpendiculaires.
C'est quand elles sont <emphasis>sécantes</emphasis> en formant un angle droit.

Dans cet <emphasis>exemple</emphasis>,
la droite <break time="100ms"/>D<break time="200ms"/> est <emphasis>perpendiculaire
</emphasis> à la droite <break time="100ms"/>F.
Elles se croisent <break time="50ms"/>en formant un angle droit.
On note comme ça,
pour dire que les droites sont <emphasis>perpendiculaires</emphasis>
<break time="100ms"/> entre elles.
voila, on a revu 3 définitions.
ce qu'il faut retenir,
c'est que des droites <emphasis>sécantes</emphasis> se <emphasis>croisent</emphasis>,
des droites <emphasis>parallèles</emphasis> se croisent <emphasis>jamais</emphasis>,
et des droites <emphasis>perpendiculaires</emphasis> se croisent <break time="200ms"/>
en formant un angle droit.
On va continuer avec <emphasis>3</emphasis> propriétés.

La propriété une<break time="100ms"/> sert à montrer que 2 droites sont <emphasis>parallèles</emphasis>,
et donc <break time="200ms"/>qu'elles se croisent jamais.
Si 2 droites sont <emphasis>perpendiculaires</emphasis> à une même troisième droite,
<break time="100ms"/> alors ces deux droites sont <emphasis>parallèles </emphasis> entre
elles.
Ici,
on sait que la droite D3, est <emphasis>perpendiculaire</emphasis> à la droite <break
time="100ms"/>D1, et aussi à la droite <break time="200ms"/>D2.
donc, D1, et, D2, sont <emphasis>parallèles</emphasis> entre elles.

Il y a une autre façon de montrer que 2 droites sont <emphasis>parallèles</emphasis>, c'est
la propriété 2.
Si 2 droites sont <emphasis>parallèles</emphasis> à une <emphasis>même </emphasis>
troisième droite, <break time="200ms"/> alors elles sont <emphasis>parallèles </emphasis>
entre elles.
Ici,
on sait que la droite D1, est <emphasis>parallèle</emphasis> à la droite <break time="100ms"/>D2,
et on sait aussi que la droite D<emphasis>3</emphasis> est <emphasis>paralèle</emphasis>
à la droite D2.
Du coup, on peut conclure grâce à la propriété 2 que D1 et D3 sont parallèles, car ce sont 2
droites qui sont parallèles à une même troisième droite.

la dernière <emphasis>propriété</emphasis> sert à montrer que 2 droites sont perpendicu-
laires.
c'est quand 2 droites se croisent en formant un angle droit.
La propriété 3, est que si 2 droites sont parallèles, et si une troisième droite est <empha-

sis>perpendiculaire</emphasis> à l'une d'elle, alors cette troisième droite est perpendiculaire à l'autre.
Ici,
on sait que les droites D1 et D2 sont parallèles, et on sait que D3 est perpendiculaire à D2.
Du coup, on peut dire que D3 est perpendiculaire à D1.

J'ai fini le cours.
Voici quelques conseils pour te faciliter la tâche.
Déjà, Il faut bien lire l'énoncé pour voir de quelle droite on parle.
Et aussi pour voir toutes les informations qu'on a.
Et enfin, quand on voit une figure, il faut regarder à quel figure-clé elle ressemble, parmi les 3 qu'on a vu.
Je les remets ici.
Il faut Utiliser la <emphasis>propriété</emphasis> qui correspond à une de ces figures <emphasis>clés</emphasis>, pour répondre aux questions.

Leçon de 5e

On va <emphasis>comencer</emphasis> par le vocabulaire.
Il y a <emphasis>4</emphasis> opérations.
L'addition, la soustraction, la division, et la multiplication.
Le résultat de <emphasis>l'addition</emphasis> s'appelle la somme.
Et le résultat de la <emphasis>soustraction</emphasis> s'appelle la différence.
Pour une division, le résultat s'appelle le quotient,
et pour une multiplication, le résultat s'appelle le produit.

Pour savoir comment <emphasis>calculer</emphasis> une expression, il faut savoir par quelle <emphasis>opération</emphasis> il faut commencer.
On va d'abord voir les <emphasis>expressions</emphasis> <emphasis>sans </emphasis> parenthèse. Il faut calculer de la <emphasis>gauche</emphasis> vers la droite.
Mais, il y a des opération qu'il faut faire en premier !
On calcule d'abord la multiplication ou la division, puis l'addition ou la soustraction.
Par exemple,
on part de la gauche, et on calcule d'abord la multiplication.
On a donc, A égal à 3+18.
On a plus qu'à <emphasis>additionner</emphasis> 3 et 18 pour avoir le résultat.
Donc A, est égal à 21.

Quand il y a des parenthèses,
Il faut <emphasis>d'abord</emphasis> calculer ce qu'il y a dans les parenthèses les plus intérieures,
tout en respectant les ordres des opérations.
d'abord la multiplication ou la division, ensuite l'addition ou la soustraction.
Par exemple,
dans cette expression,
on calcule <emphasis>d'abord</emphasis> ce qu'il y a dans la <emphasis>parenthèse</emphasis> la plus intérieure.

On a donc entre parenthèse 3+7, multiplié par 5.
Il nous reste une parenthèse. On calcule donc ce qu'il y a dedans.
Il nous reste plus qu'à <emphasis>calculer<emphasis> cette multiplication.
Et on a, B, égal a 50.

On a vu le <emphasis>vocabulaire</emphasis> des <emphasis>résultats</emphasis> des opérations, avec la somme pour l'addition, la différence pour la soustraction, le quotient pour la division, et le produit pour la multiplication.
Maintenant, il faut savoir si une expression est une somme, une différence, un quotient, ou un produit.
Et pour ça, il faut voir quelle est la dernière opération qu'il faut calculer. On va voir 2 exemples.
Ici, on a A, égal à 3+4 multiplié par 5.
Il faut <emphasis>d'abord</emphasis> calculer la multiplication 4 fois 5, parce qu'on commence par la multiplication.
Donc, A égale à 3 plus 20.
Pour finir cette <emphasis>expression</emphasis>, on fait une addition.
Donc, A, est une somme, parce qu'on a <emphasis>fini</emphasis> par une addition.

Maintenant, on va voir un exemple avec des parenthèses.
Quand il y a des parenthèses, il faut commencer par là.
Ici, on a, B, égal à <emphasis>20<emphasis> divisé par 7 moins 2, entre parenthèses.
Comme il faut commencer par ce qu'il y a dans les parenthèses, on fait <emphasis>d'abord</emphasis> 7 moins 2.
Donc, B, égal à 20 divisé par 5.
On fini par faire une <emphasis>division</emphasis>, du coup, B, est un quotient.

J'ai fini le cours.
Voici quelques conseils pour te faciliter la tâche.
Pour <emphasis>calculer</emphasis> une expression, il faut voir ce qu'il faut calculer en premier.
On calcule <emphasis>d'abord</emphasis> ce qu'il y a dans les parenthèses, en appliquant les règles des <emphasis>priorités</emphasis> opératoires.
D'abord la multiplication ou la division, ensuite, l'addition ou la soustraction.
Et il faut calculer de gauche à droite.
Par contre, quand on ne demande <emphasis>pas</emphasis> le résultat de l'expression, et qu'on demande si c'est une somme, ou une différence, ou un quotient, ou un produit, il faut voir la <emphasis>dernière</emphasis> opération à calculer.

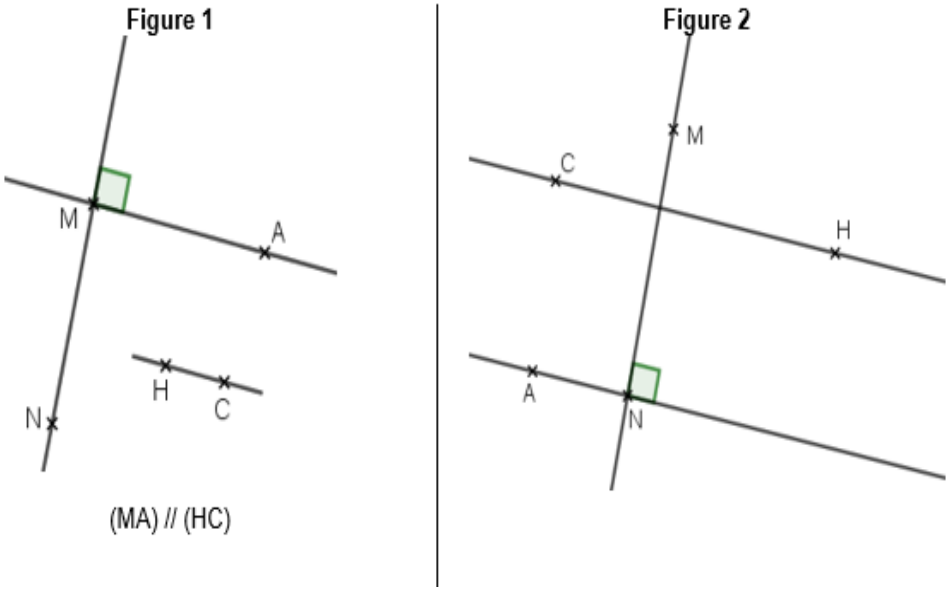
Fin du discours

Voilà, je vais te laisser faire le test de maths.

F. Test standardisé de mathématiques 1 pour les 6e.

Exercice 1

1/ Indiquer pour quelle figure on est certain que les droites (MN) et (HC) sont perpendiculaires. Cliquer sur la réponse pour la sélectionner.



- Figure 1.
- Figure 2.
- Je ne sais pas.

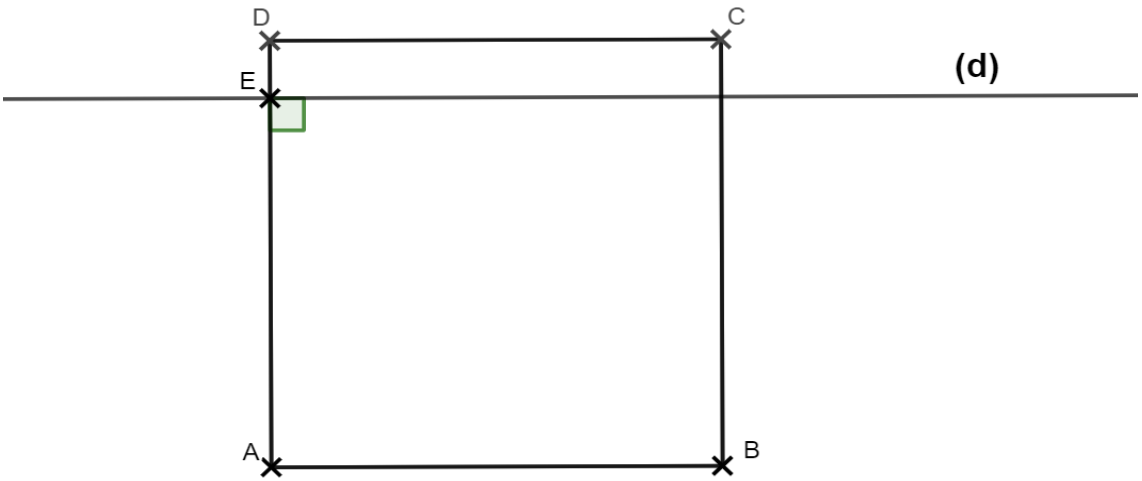
Label : t1_exo1_A.
Barème : « figure 1 » : 3 points.

b/ Choisir la propriété du cours qui permet de justifier la réponse. Cliquer sur la réponse pour la sélectionner.

- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Label : t1_exo1_B.
Barème : « propriété 3 » : 3 points.

Exercice 2



ABCD est un carré, E est un point de [AD] et la droite (d) est la perpendiculaire à la droite (AD) passant par E.

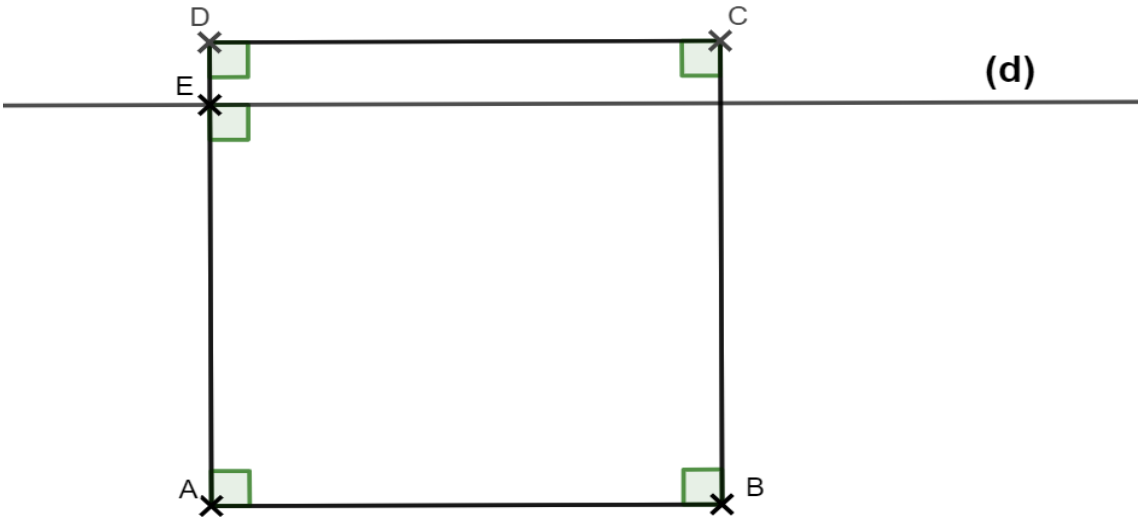
1/ a/ Que peut-on dire des droites (AB) et (AD) ?
On peut dire que les droites (AB) et (AD) sont (liste déroulante) : confondues / sécantes / parallèles / perpendiculaires / je ne sais pas.

Label : t1_exo2_1.
Barème : « sécantes » : 0.5 point. « perpendiculaires » : 1 point.

b/ Pourquoi ?

- Parce que je le vois.
- Parce que (d) est perpendiculaire à (AD).
- Parce que ABCD est un quadrilatère.
- Parce que ABCD est un rectangle.
- Parce que ABCD est un carré.
- Parce que ABCD est un losange.
- Parce que j'ai vérifié avec mon équerre.
- Je ne sais pas.

Exercice rajouté pour l'étude principale.
Barème : « Parce que ABCD est un carré » : 2 points. « Parce que ABCD est un rectangle » : 1 point. « Parce que j'ai vérifié avec mon équerre » : 0.5 point.



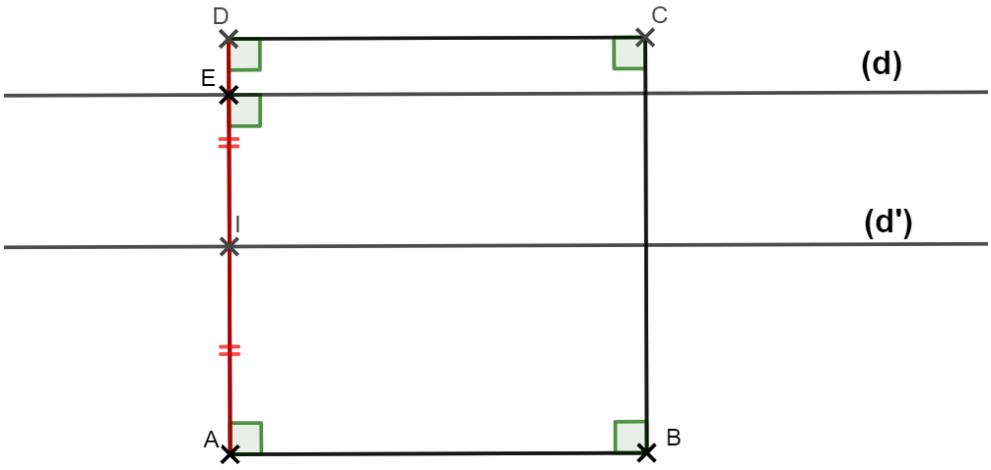
2/ Que peut-on dire des droites (AB) et (d) ?
On peut dire que les droites (AB) et (d) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.
Label : t1_exo2_2.
Barème : « parallèle » : 1 point.

3/ Compléter la propriété qui permet de justifier la réponse à la question 2.

Si deux *(liste déroulante 1)* sont *(liste déroulante 2) (liste déroulante 3)*, alors elles sont *(liste déroulante 4)* entre elles.

- *Liste déroulante 1* : droites / segments / demi-droites / cercles.
- *Liste déroulante 2* : parallèles / sécantes / perpendiculaires / confondues.
- *Liste déroulante 3* : à une même troisième droite / et si une troisième droite est perpendiculaire à l'une d'elles / et si une troisième droite est parallèle à l'une d'elles.
- *Liste déroulante 4* : parallèles / sécantes / perpendiculaires / confondues.

Label : t1_exo2_3.
Barème :
- *Phrase correcte (4 choix corrects : droites, perpendiculaires, à une même troisième droite, parallèles) : 3 points.*
- *Première partie de la phrase correcte « Si deux droites sont perpendiculaires à une même troisième droite » mais pas la deuxième partie (3 premiers choix corrects : droites, perpendiculaires, à une même troisième droite) : 1 point.*
- *Deuxième partie de la phrase correcte (quatrième choix : « parallèles ») ou en cohérence avec la réponse à la question 2 (si l'élève a répondu « perpendiculaires » à la question 2 et choisit « perpendiculaires » au quatrième choix de la question 3) : 1 point.*



On trace la droite (d') parallèle à la droite (AB) passant par le milieu I de [AD].

4/ Que peut-on dire des droites (d') et (d) ?
On peut dire que les droites (d') et (d) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.

Label : t1_exo2_4.
Barème : « parallèles » : 1 point.

5/ Quelle est la propriété qui permet de justifier la réponse à la question 4 ?

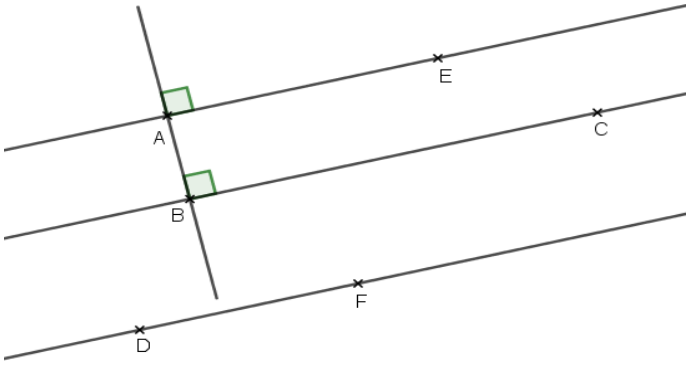
- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Label : t1_exo2_5.
Barème : « Propriété 2 » : 3 points. « Propriété 1 » : 1 point.

G. Test standardisé de mathématiques 2 pour les 6e.

Exercice 1

On considère la figure ci-dessous dans laquelle les droites (BC) et (DF) sont parallèles.



1/ a/ Compléter les phrases suivantes :

Les droites (AE) et (AB) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.
Label : t2_exo1_A_reponseA.
Barème : « perpendiculaires » : 1 point et « sécantes » : 0.5 point.

Les droites (BC) et (AB) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.
Label : t2_exo1_A_reponseB.
Barème : « perpendiculaires » : 1 point et « sécantes » : 0.5 point.

Les exercices 1/ b/ et 1/ c/ ci-dessous sont uniquement dans les prétests (Chapitre 8).

b/ Que peut-on en déduire pour les droites (AE) et (BC) ?
On peut en déduire que les droites (AE) et (BC) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.

Label : t2_exo1_B.
Barème : « parallèles » : 2 points

c/ Choisir la propriété du cours qui permet de justifier la réponse à la question 1/ b/.

- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.

- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Label : t2_exo1_C.
Barème : « Propriété 1 » : 3 points.

L'exercice 1/ b/ ci-dessous est uniquement dans l'étude principale (Chapitre 9)

b/ Choisir la propriété du cours qui permet de justifier ceci : les droites (AE) et (BC) sont parallèles.

- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Barème : « Propriété 1 » : 3 points.

2/ a/ Compléter avec le symbole qui convient :

(AB) *(liste déroulante)* (BC)

Label : t2_exo1_2_A_reponseA.
Barème : « ⊥ » : 1 point.

(BC) *(liste déroulante)* (DF)

Label : t2_exo1_2_A_reponseB.
Barème : « // » : 1 point.

Liste déroulante : « // »; « ⊥ »; « ∈ »; « je ne sais pas »

b/ Que peut-on en déduire pour les droites (AB) et (DF) ?

On peut en déduire que les droites (AB) et (DF) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.

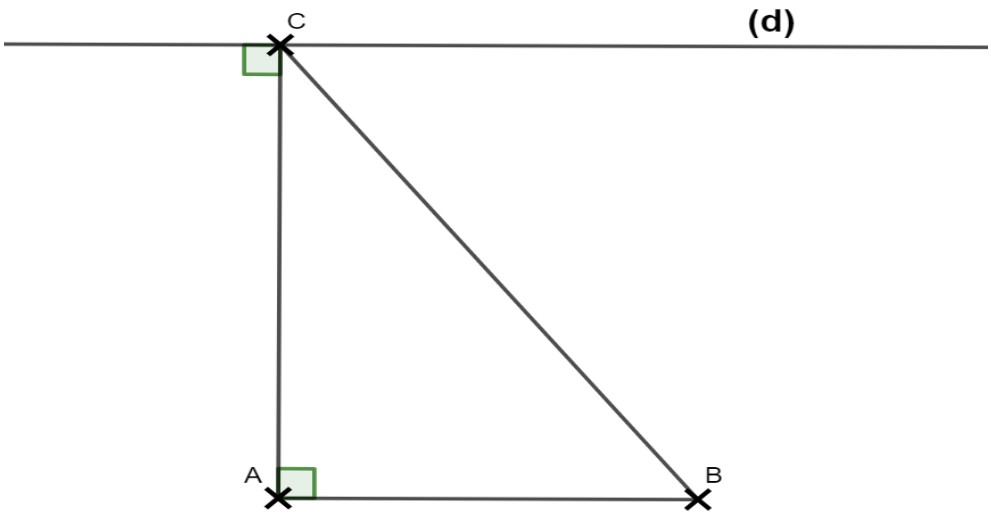
Label : t2_exo1_2_B.
Barème : « perpendiculaires » : 2 points.

c/ Choisir la propriété du cours qui permet de justifier la réponse à la question 2/ b/.

- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Label : t2_exo1_2_C.
Barème : « Propriété 3 » : 3 points.

Exercice 2



Le triangle ABC est rectangle en A et (d) est la perpendiculaire à (AC) passant par C.

1/ a/ Que peut-on dire des droites (AB) et (AC) ?

Les droites (AB) et (AC) sont (liste déroulante) : sécantes / confondues / parallèles / perpendiculaires.

Label : t2_exo2_1.
Barème : « perpendiculaires » : 1 point et « sécantes » : 0.5 point.

b/ Pourquoi ?

- Parce que je le vois.
- Parce que (d) est perpendiculaire à (AC).
- Parce que ABC est un triangle.

- Parce que ABC est un triangle isocèle.
- Parce que ABC est un triangle équilatéral.
- Parce que ABC est un triangle rectangle.
- Parce que j'ai vérifié avec mon équerre.
- Je ne sais pas.

Exercice rajouté pour l'étude principale.
Barème : « Parce que ABC est un triangle rectangle » : 2 points. « Parce que j'ai vérifié avec mon équerre » : 0.5 point.

2/ Que peut-on dire des droites (AB) et (d) ?

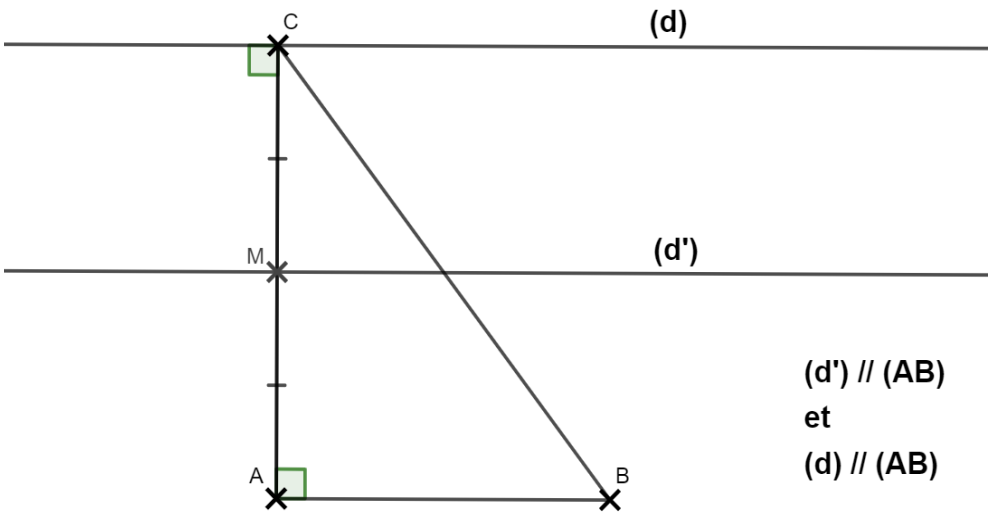
Les droites (AB) et (d) sont sécantes / parallèles / perpendiculaires / confondues (liste déroulante).

Exercice rajouté pour l'étude principale.
Barème : « parallèles » : 1 point.

3/ Choisir la propriété du cours qui permet de justifier la réponse à la question 2. Choisis la propriété uniquement si tu es sûr de ta réponse, choisis "je ne sais pas" sinon.

- Propriété 1 : Si deux droites sont perpendiculaires à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 2 : Si deux droites sont parallèles à une même troisième droite, alors elles sont parallèles entre elles.
- Propriété 3 : Si deux droites sont parallèles et si une troisième droite est perpendiculaire à l'une d'elles, alors cette troisième droite est perpendiculaire à l'autre.
- Je ne sais pas.

Label : t2_exo2_2.
Exercice modifié pour l'étude principale.
Barème : « Propriété 1 » : 3 points.



On trace la droite (d') parallèle à (AB) passant par le milieu M de [AC].

4/ Que peut-on dire des droites (d') et (d) ?

Les droites (d') et (d) sont (liste déroulante) : sécantes / parallèles / perpendiculaires / confondues.

Label : t2_exo2_3.
Barème : « parallèles » : 1 point.

5/ Compléter la propriété qui permet de justifier la réponse à la question 4.

Si deux *(liste déroulante 1)* sont *(liste déroulante 2) (liste déroulante 3)*, alors elles sont *(liste déroulante 4)* entre elles.

- *Liste déroulante 1* : droites / segments / demi-droites / cercles.
- *Liste déroulante 2* : parallèles / sécantes / perpendiculaires / confondues.
- *Liste déroulante 3* : à une même troisième droite / et si une troisième droite est perpendiculaire à l'une d'elles / et si une troisième droite est parallèle à l'une d'elles.
- *Liste déroulante 4* : parallèles / sécantes / perpendiculaires / confondues.

Label : t2_exo2_4.
Barème :

- *Phrase correcte (4 choix corrects : droites, parallèles, à une même troisième droite, parallèles) : 3 points.*
- *Première partie de la phrase correcte « Si deux droites sont parallèles à une même troisième droite » mais pas la deuxième partie (3 premiers choix corrects : droites, parallèles, à une même troisième droite) : 1 point.*

- *Deuxième partie de la phrase correcte (quatrième choix : « parallèles ») ou en cohérence avec la réponse à la question 2 (si l'élève a répondu « perpendiculaires » à la question 2 et choisit « perpendiculaires » au quatrième choix de la question 3) : 1 point.*

H. Test standardisé de mathématiques 1 pour les 5e.

Exercice 1

1/ Compléter les expressions suivantes à l'aide des deux opérateurs donnés pour que les affirmations soient vraies.

Affirmations	Opérateurs	Expressions
A est une somme. *Label : t1_exo1_1_premiereLigne.* *Correction : (8-4)+(3x6).*	+ et x	A=(8-4)...(3...6)
B est un quotient. *Label : t1_exo1_1_deuxiemeLigne.* *Correction : 36÷(6+15).*	+ et ÷	B=36...(6...15)
C est une différence. *Label : t1_exo1_1_troisiemeLigne.* *Correction : (28÷(3+4))-(8x6).*	÷et -	C=(28...(3+4))...(8x6)

Barème : 1 point par expression correcte.

2/ Pour chacune des expressions suivantes préciser quelle opération sera effectuée en dernier, compléter l'affirmation pour qu'elle soit vraie et donner le résultat de l'expression.

Expressions	Dernière opération effectuée	Affirmation	Résultat
D=10-4x2	(liste déroulante) • - • x	D est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t1_exo2_1_premiereLigne. *Barème : -, une différence, 2 : 1 point.*			
E=(25-5)÷4+1	(liste déroulante) • - • ÷ • +	E est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t1_exo2_1_deuxiemeLigne. *Barème :* • +, *une somme, 20 : 0.5 point.* • +, *une somme, 4 : 0.5 point.* • +, *une somme, 6 : 1.5 points.*			
$F = \frac{2x5}{3+7}$	(liste déroulante) • x • ÷ • +	F est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t1_exo2_1_troisiemeLigne. *Barème :* • ÷, *un quotient, 0.7 : 0.5 point.* • ÷, *un quotient, 1 : 1.5 points.*			

Exercice 2

1/ Traduire les expressions numériques suivantes à l'aide d'une phrase.
Exemple : 10 + 2 × 5 : La somme de dix et du produit de 2 par 5.

a/ 25 – 6 × 4

(liste déroulante 1) de *(liste déroulante 2)* et *(liste déroulante 3)* de *(liste déroulante 4)* par *(liste déroulante 5)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : 25 / 6 / 4.
- *Liste déroulante 3* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 4* : 25 / 6 / 4.
- *Liste déroulante 5* : 25 / 6 / 4.

Label : t1_exo2_1_a.
Barème :
- *Phrase correcte (La différence de 25 et du produit de 6 par 4) : 1.5 points.*
- *Première opération correcte « La différence » (identification de la nature de l'expression) : 0.5 point.*

b/ (15 + 25) ÷7

(liste déroulante 1) de *(liste déroulante 2)* et *(liste déroulante 3)* de *(liste déroulante 4)* par *(liste déroulante 5)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : 15 / 25 / 7.
- *Liste déroulante 4* : 15 / 25 / 7.
- *Liste déroulante 5* : 15 / 25 / 7.

Label : t1_exo2_1_b.
Barème :
- *Phrase correcte (Le quotient de la somme de 15 et de 25 par 7) : 1.5 points.*
- *Première opération correcte « Le quotient » (identification de la nature de l'expression) : 0.5 point.*

c/ (10 + 5) × (7 – 2)

(liste déroulante 1) (liste déroulante 2) de *(liste déroulante 3)* et de *(liste déroulante 4)* par *(liste déroulante 5)* de *(liste déroulante 6)* et de *(liste déroulante 7)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : 10 / 5 / 7 / 2.

- *Liste déroulante 4* : 10 / 5 / 7 / 2.
- *Liste déroulante 5* : la différence / la somme / le produit / le quotient.
- *Liste déroulante 6* : 10 / 5 / 7 / 2.
- *Liste déroulante 7* : 10 / 5 / 7 / 2.

Label : t1_exo2_1_c.
Barème :

- *Phrase correcte (Le produit de la somme de 10 et de 5 par la différence de 7 et de 2) : 2 points.*
- *Première opération correcte « Le produit » (identification de la nature de l'expression) : 1 point.*

d/ $5 \times 7 + 2 \times 3$

(liste déroulante 1) (liste déroulante 2) de *(liste déroulante 3)* par *(liste déroulante 4)* et *(liste déroulante 5)* de *(liste déroulante 6)* par *(liste déroulante 7)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : 5 / 7 / 2 / 3.
- *Liste déroulante 4* : 5 / 7 / 2 / 3.
- *Liste déroulante 5* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 6* : 5 / 7 / 2 / 3.
- *Liste déroulante 7* : 5 / 7 / 2 / 3.

Label : t1_exo2_1_d.
Barème :

- *Phrase correcte (La somme du produit de 5 par 7 et du produit de 2 par 3) : 2 points.*
- *Première opération correcte « La somme » (identification de la nature de l'expression) : 1 point.*

e/ $12 + \frac{10}{2}$

(liste déroulante 1) de *(liste déroulante 2)* et *(liste déroulante 3)* de *(liste déroulante 4)* par *(liste déroulante 5)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : 12 / 10 / 2.
- *Liste déroulante 3* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 4* : 12 / 10 / 2.
- *Liste déroulante 5* : 12 / 10 / 2.

Label : t1_exo2_1_e.
Barème :

- *Phrase correcte (La somme de 12 et du quotient de 10 par 2) : 2 points.*
- *Première opération correcte « La somme » (identification de la nature de l'expression) : 1*

point.

f/ $\frac{25-5}{10}$

(liste déroulante 1) (liste déroulante 2) de *(liste déroulante 3)* et de *(liste déroulante 4)* par *(liste déroulante 5).*

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : 25 / 5 / 10.
- *Liste déroulante 4* : 25 / 5 / 10.
- *Liste déroulante 5* : 25 / 5 / 10.

Label : t1_exo2_1_f.
Barème :

- *Phrase correcte (Le quotient de la différence de 25 et de 5 par 10) : 2 points.*
- *Première opération correcte « Le quotient » (identification de la nature de l'expression) : 1 point.*

2/ Écrire les expressions numériques qui correspondent aux phrases suivantes. Exemple : La somme de 8 et du quotient de 10 par 2 : 8 + 10 ÷2.

a/ La somme de 5 et du produit de 3 par 4.

Label : t1_exo2_2_a.
Exercice supprimé de l'étude principale.
Correction : 5+3×4 ou 3×4+5 ou 4×3+5.
Barème : expression correcte : 1.5 points.

b/ La différence entre le produit de 3 par 8 et 5.
Label : t1_exo2_2_b.
Exercice supprimé de l'étude principale.
Correction : 3 × 8 – 5 ou 8 × 3 - 5.
Barème : expression correcte : 1.5 points.

c/ Le produit de la somme de 4 et de 3 par le quotient de 12 par 2.
Label : t1_exo2_2_c.
Correction : (4+3)x(12÷2) ou les mêmes expressions avec commutation (3+4)×(12÷2) ou (12÷2)×(4+3) ou (12÷2)×(3+4) ou les mêmes sans parenthèses autour de 12÷2.
Barème : expression correcte : 2 points. Expression sans parenthèse autour de 4+3 : 0.5 point.

d/ La somme de 25 et du quotient de 10 par 5.
Label : t1_exo2_2_d.
Correction : 25+10÷5 ou 10÷5+25 ou 25+(10÷5) ou (10÷5)+25.
Barème : expression correcte : 1.5 points.

e/ Le quotient de la somme de 4 et 6 par le produit de 5 par 4.
Label : t1_exo2_2_e.
Correction : (4+6)÷(5×4) ou (6+4)÷(5×4) ou (4+6)÷(4×5) ou (6+4)÷(4×5).
Barème : expression correcte : 2 points. Expression sans parenthèse : 0.5 point.

I. Test standardisé de mathématiques 2 pour les 5e.

Exercice 1

1/ Compléter les expressions suivantes à l'aide des deux opérateurs donnés pour que les affirmations soient vraies.

Affirmations	Opérateurs	Expressions
A est un produit. *Label : t2_exo1_1_premiereLigne.* *Correction : (11-4)x(4+5).*	+ et x	A=(11-4)...(4...5)
B est un quotient. *Label : t2_exo1_1_deuxiemeLigne.* *Correction : 45÷(24-9).*	- et ÷	B=45...(24...9)
C est une différence. *Label : t2_exo1_1_troisiemeLigne.* *Correction : (28÷(3+4))-(8x6).*	÷et -	C=(28...(3+4))...(8x6)

Barème : 1 point par expression correcte.

2/ Pour chacune des expressions suivantes préciser quelle opération sera effectuée en dernier, compléter l'affirmation pour qu'elle soit vraie et donner le résultat de l'expression.

Expressions	Dernière opération effectuée	Affirmation	Résultat
D=16-6x2	(liste déroulante) • - • x	D est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t2_exo2_1_premiereLigne. *Barème : -, une différence, 4 : 1 point.*			
E=(45-9)÷4+5x3	(liste déroulante) • + • - • ÷ • x	E est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t2_exo2_1_deuxiemeLigne. *Barème :* • +, *une somme, 42 : 0.5 point.* • +, *une somme, 24 : 1.5 points.*			
$F = \frac{26-2x5}{2+3x2}$	(liste déroulante) • - • + • x • ÷	F est (liste déroulante) • une somme • une différence • un produit • un quotient	*saisie du résultat*
Label : t2_exo2_1_troisiemeLigne. *Barème :* • ÷, *un quotient, 15 : 0.5 point.* • ÷, *un quotient, 1.6 : 0.5 point.* • ÷, *un quotient, 2 : 2 points.*			

Exercice 2

1/ Traduire les expressions numériques suivantes à l'aide d'une phrase.
Exemple : 12 + 3 × 4 : La somme de 12 et du produit de 3 par 4.

a/ 40 - 7 × 3

(liste déroulante 1) de *(liste déroulante 2)* et *(liste déroulante 3)* de *(liste déroulante 4)* par *(liste déroulante 5)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : 40 / 7 / 3.
- *Liste déroulante 3* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 4* : 40 / 7 / 3.
- *Liste déroulante 5* : 40 / 7 / 3.

Label : t2_exo2_1_a.
Barème :
- *Phrase correcte (La différence de 40 et du produit de 7 par 3) : 1.5 points.*
- *Première opération correcte « La différence » (identification de la nature de l'expression) : 0.5 point.*

b/ 35 ÷(5+2)

(liste déroulante 1) de *(liste déroulante 2)* par *(liste déroulante 3)* de *(liste déroulante 4)* et de *(liste déroulante 5)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : 35 / 5 / 2.
- *Liste déroulante 3* : la différence / la somme / le produit / le quotient.
- *Liste déroulante 4* : 35 / 5 / 2.
- *Liste déroulante 5* : 35 / 5 / 2.

Label : t2_exo2_1_b.
Barème :
- *Phrase correcte (Le quotient de 35 par la somme de 5 et de 2) : 1.5 points.*
- *Première opération correcte « Le quotient » (identification de la nature de l'expression) : 0.5 point.*

c/ (13 + 7) × (12 – 2)

(liste déroulante 1) (liste déroulante 2) de *(liste déroulante 3)* et de *(liste déroulante 4)* par *(liste déroulante 5)* de *(liste déroulante 6)* et de *(liste déroulante 7)*.

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : 13 / 7 / 12 / 2.

- *Liste déroulante 4* : 13 / 7 / 12 / 2.
- *Liste déroulante 5* : la différence / la somme / le produit / le quotient.
- *Liste déroulante 6* : 13 / 7 / 12 / 2.
- *Liste déroulante 7* : 13 / 7 / 12 / 2.

Label : t2_exo2_1_c.
Barème :

- *Phrase correcte (Le produit de la somme de 13 et de 7 par la différence de 12 et de 2) : 2 points.*
- *Première opération correcte « Le produit » (identification de la nature de l'expression) : 0.5 point.*

d/ 15 + $\frac{12}{3}$

(liste déroulante 1) de *(liste déroulante 2)* et *(liste déroulante 3)* de *(liste déroulante 4)* par *(liste déroulante 5).*

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : 15 / 12 / 3.
- *Liste déroulante 3* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 4* : 15 / 12 / 3.
- *Liste déroulante 5* : 15 / 12 / 3.

Label : t2_exo2_1_d.
Barème :

- *Phrase correcte (La différence de 15 et du quotient de 12 par 3) : 1.5 points.*
- *Première opération correcte « La différence » (identification de la nature de l'expression) : 0.5 point.*

e/ $\frac{40-4}{7+11}$ **+ 5x3**

(liste déroulante 1) (liste déroulante 2) (liste déroulante 3) de *(liste déroulante 4)* et de *(liste déroulante 5)* par *(Liste déroulante 6)* de *(Liste déroulante 7)* et de *(Liste déroulante 8)* et *(Liste déroulante 9)* de *(Liste déroulante 10)* par *(Liste déroulante 11).*

- *Liste déroulante 1* : La différence / La somme / Le produit / Le quotient.
- *Liste déroulante 2* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 3* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 4* : 40 / 4 / 7 / 11 / 5 / 3.
- *Liste déroulante 5* : 40 / 4 / 7 / 11 / 5 / 3.
- *Liste déroulante 6* : la différence / la somme / le produit / le quotient.
- *Liste déroulante 7* : 40 / 4 / 7 / 11 / 5 / 3.
- *Liste déroulante 8* : 40 / 4 / 7 / 11 / 5 / 3.
- *Liste déroulante 9* : de la différence / de la somme / du produit / du quotient.
- *Liste déroulante 10* : 40 / 4 / 7 / 11 / 5 / 3.
- *Liste déroulante 11* : 40 / 4 / 7 / 11 / 5 / 3.

Label : t2_exo2_1_e.
Barème :

- *Phrase correcte (La somme du quotient de la différence de 40 et de 4 par la somme de 7 et de 11 et du produit de 5 par 3) : 2 points.*
- *Première opération correcte « La somme » (identification de la nature de l'expression) : 0.5 point.*

2/ Écrire les expressions numériques qui correspondent aux phrases suivantes.
Exemple : La somme de 8 et du quotient de 10 par 2 : 8 + 10 ÷2.
a/ La différence de 25 et du produit de 3 par 4.

Label : t2_exo2_2_a.
Exercice supprimé de l'étude principale.
Correction : 25 – 3 × 4 ou 25 – 4 × 3.
Barème : expression correcte : 1.5 points.

b/ La somme du quotient de 15 par 3 et de 11.
Label : t2_exo2_2_b.
Exercice supprimé de l'étude principale.
Correction : (15 ÷3) + 11 ou 15 ÷3 + 11 ou 11 + 15 ÷3 ou 11 + (15 ÷3).
Barème : expression correcte : 1.5 points.

c/ Le produit de la somme de 7 et de 2 par le quotient de 20 par 4.
Label : t2_exo2_2_c.
Correction : (7 + 2) × (20 ÷4) ou (7 + 2) × 20 ÷4 ou (2 + 7) × (20 ÷4) ou (2 + 7) × 20 ÷4.
Barème : expression correcte : 2 points. Expression sans parenthèse autour de 7+2 : 0.5 point.

d/ La somme du produit de 5 par 4 et du quotient de 28 par 7.
Label : t2_exo2_2_d.
Correction : (5 × 4) + (28 ÷7) ou 5 × 4 + 28 ÷7 ou (4 × 5) + (28 ÷7) ou 4 × 5 + 28 ÷7.
Barème : expression correcte : 2 points.

e/ Le quotient de la différence de 34 et de 20 par la somme de 6 et du produit de 3 par 4.
Label : t2_exo2_2_e.
Correction : (34 – 20) ÷(6 + 3 × 4) ou (34 – 20) ÷(3 × 4 + 6).
Barème : expression correcte : 2 points. Expression sans parenthèse autour de 34-20 ou autour de 6+3x4 : 0.5 point.

J. Questionnaire d'évaluation des agents virtuels pédagogiques (études pilotes 1A et 1B du Chapitre 7.)

TABLEAU 1

Items utilisés pour la mesure de la compétence en mathématiques perçue de l'agent virtuel pédagogique.

1. Est-ce que tu penses que le lycéen/la lycéenne en 3D réussit bien en maths ?
2. Est-ce que tu penses que le lycéen/la lycéenne en 3D travaille et révise ses leçons de maths ?
3. Est-ce que tu penses que le lycéen/la lycéenne en 3D passe du temps à faire ses devoirs en maths ?

TABLEAU 2

Items utilisés pour la mesure du niveau d'anthropomorphisme perçu de l'agent virtuel pédagogique, adaptés du questionnaire du Godspeed évaluant l'anthropomorphisme et celle évaluant l'animation de BARTNECK *et al. (2009).*

Item original	Item adapté et utilisé
Échelle évaluant l'anthropomorphisme	
1. Machinelike - Humanlike	J'ai trouvé que le lycéen/la lycéenne en 3D avait un aspect robotique.
2. Machinelike - Humanlike	J'ai trouvé que le lycéen/la lycéenne en 3D ressemblait à un humain.
3. Moving rigidly - Moving elegantly	J'ai trouvé que le lycéen/la lycéenne en 3D avait du mal à bouger.
4. Fake - Natural	J'ai trouvé que le lycéen/la lycéenne en 3D était réaliste.
5. Fake - Natural	J'ai trouvé que l'apparence du lycéen/ de la lycéenne en 3D était naturelle.
Échelle évaluant l'animation	
6. Artificial - Lifelike	J'ai trouvé que le lycéen/la lycéenne en 3D bougeait comme un robot.
7. Mechanical - Organic	J'ai trouvé que le lycéen/la lycéenne en 3D bougeait comme un humain.

Note. Le questionnaire du Godspeed étant une échelle d'évaluation, nous avons adapté certains éléments en deux items sur une échelle de Likert.

TABLEAU 3

Items utilisés pour la mesure de la sympathie perçue de l'agent virtuel pédagogique, adaptés du questionnaire du Godspeed de BARTNECK *et al. (2009).*

Item original	Item adapté et utilisé
1. Unfriendly - Friendly	J'ai trouvé que le lycéen/la lycéenne en 3D était amical.
2. Unkind - Kind	J'ai trouvé que le lycéen/la lycéenne en 3D était gentil(le).
3. Unkind - Kind	J'ai trouvé que le lycéen/la lycéenne en 3D était méchant(e).
4. Awful - Nice	J'ai trouvé que le lycéen/la lycéenne en 3D était sympa.
5. Unpleasant - Pleasant	J'ai trouvé que le lycéen/la lycéenne en 3D était désagréable.

Note. Le questionnaire du Godspeed étant une échelle d'évaluation, nous avons adapté certains éléments en deux items sur une échelle de Likert.

TABLEAU 4

Items utilisés pour la mesure de l'identification à l'agent virtuel pédagogique, adaptés de O'BRIEN *et al. (2017).*

Item original	Item adapté et utilisé
1. How much are you similar to the workshop leader you chose?	Penses-tu avoir des points communs avec le lycéen/la lycéenne en 3D?
2. How much do you look up to the workshop leader you chose?	Est-ce que tu voudrais avoir plus de points communs avec le lycéen/la lycéenne en 3D?
3. How much do you like the workshop leader you chose?	À quel point aimes-tu lycéen/lycéenne en 3D?
4. How much do you admire the workshop leader you chose?	À quel point trouves-tu le lycéen/la lycéenne en 3D épatant(e)?

TABLEAU 5

Items utilisés pour la mesure de l'identification à la réussite en mathématiques de l'agent virtuel pédagogique.

1. Comme le lycéen/la lycéenne en 3D, j'ai envie de faire de mon mieux en maths.
2. Comme le lycéen/la lycéenne en 3D, je considère qu'il est important pour moi de réussir en maths.
3. Comme le lycéen/la lycéenne en 3D, je suis capable d'avoir des bonnes notes en maths.
4. Comme le lycéen/la lycéenne en 3D, je pense que je pourrais mieux réussir en maths si je travaillais encore plus.

TABLEAU 6

Items utilisés pour la mesure du sentiment de rapport virtuel perçu lors de l'interaction avec l'agent virtuel pédagogique, adaptés du questionnaire du Virtual Scale Intimacy (VIS) de POTDEVIN *(2020) et du questionnaire de la perception de l'interaction de* CUPERMAN *et* ICKES *(2009).*

Item original	Item adapté et utilisé
VIS	
1. A a été spontanée avec B.	J'ai trouvé que le lycéen/la lycéenne en 3D a été naturel(le) avec moi.
2. A s'est dévoilée à B.	J'ai trouvé que le lycéen/la lycéenne en 3D s'est confié(e) à moi.
3. A a donné son avis à B.	J'ai trouvé que le lycéen/la lycéenne en 3D a donné son avis.
4. A a exprimé ses émotions à B.	J'ai trouvé que le lycéen/la lycéenne en 3D a exprimé ses émotions.
5. A a apprécié échanger avec B.	J'ai trouvé que le lycéen/la lycéenne en 3D a apprécié me parler.
6. A a été avenante avec B.	J'ai trouvé que le lycéen/la lycéenne en 3D a été gentil(le) avec moi.
7. A a cherché à se rapprocher de B.	J'ai trouvé que le lycéen/la lycéenne en 3D a cherché à être mon ami(e).
8. A a été serviable avec B.	J'ai trouvé que le lycéen/la lycéenne en 3D a voulu m'aider.
9. A a compris B.	J'ai trouvé que le lycéen/la lycéenne en 3D a compris mes attentes en maths.
10. A a compris B.	J'ai trouvé que le lycéen/la lycéenne en 3D a compris mon niveau en maths.
11. A a inspiré confiance à B.	J'ai trouvé que le lycéen/la lycéenne en 3D m'a inspiré confiance.
12. A était disposée à écouter les problèmes de B.	J'ai trouvé que le lycéen/la lycéenne en 3D était disposé(e) à m'écouter.
13. A s'est souciée de B.	J'ai trouvé que le lycéen/la lycéenne en 3D s'est soucié(e) de moi.
Perception de l'interaction	
14. I felt uncomfortable during the interaction.	Je me suis senti(e) mal à l'aise avec le lycéen/la lycéenne en 3D.
15. The interaction with the character was pleasant and interesting.	Ce que disait le lycéen/la lycéenne en 3D était cool et intéressant.
16. I would like to interact more with the character in the future.	J'aimerais revoir le lycéen/la lycéenne en 3D une autre fois.

K. TAI genre-mathématiques-lecture passé par les élèves dans l'étude principale (chapitre 9)

Code disponible en open-source :
https://github.com/ArmandoMarjorie/French_child_IAT

Mots par catégories

- Maths
  - Addition
  - Calcul
  - Droite
  - Maths
  - Nombre
- Lecture
  - Lecture
  - Lettre
  - Lire
  - Livre
  - Roman
- Garçon
  - Djibril
  - Hugo
  - Léo
  - Lucas
  - Mohamed
- Fille
  - Aminata
  - Emma
  - Nadia
  - Sarah
  - Zoë

Instructions sur ordinateur

FIGURE 1.

Capture d'écran des instructions globales pour effectuer le TAI sur ordinateur.

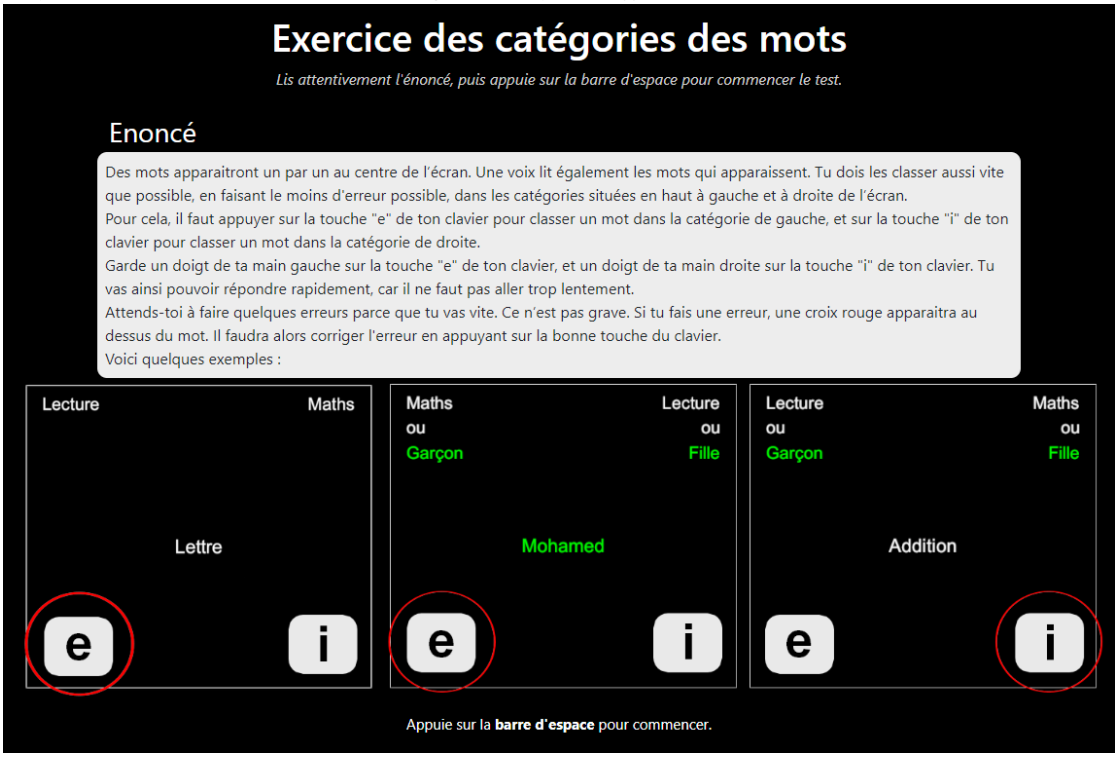
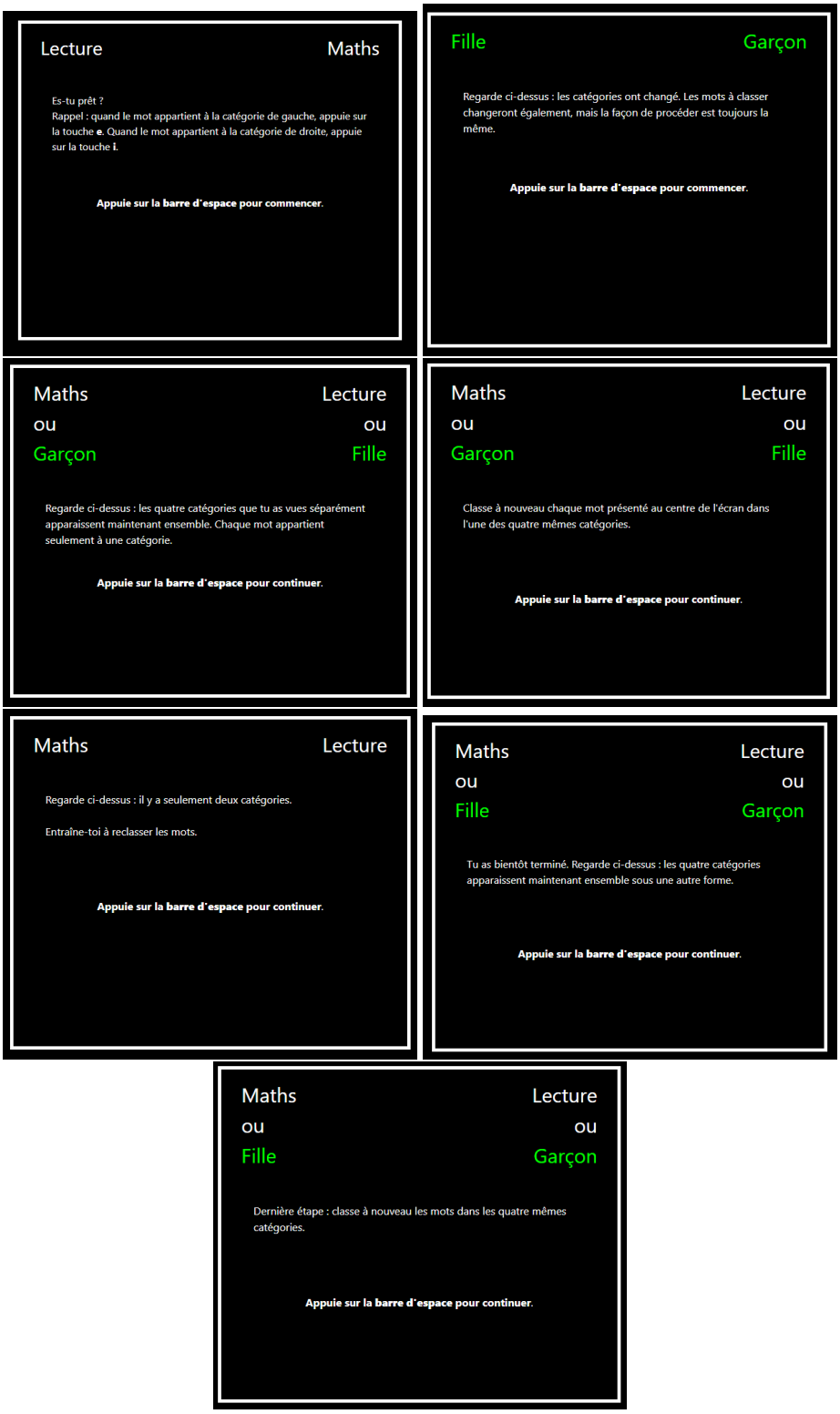


FIGURE 2.

Captures d'écran des instructions de chaque block pour effectuer le TAI sur ordinateur.



Instructions sur tablette

FIGURE 3.

Capture d'écran des instructions globales pour effectuer le TAI sur tablette.



FIGURE 4.

Captures d'écran des instructions de chaque block pour effectuer le TAI sur tablette.

