

DOCTORAT AIX-MARSEILLE UNIVERSITÉ
délivré par
Université de Provence

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR D'AIX-MARSEILLE UNIVERSITÉ

**Formation doctorale :
Psychologie**

**Présentée et soutenue publiquement
par**

Jean DELPECH

le 17 décembre 2010

**Stéréotypes sociaux et régulation de la différence de sexe
en matière de rotations mentales 3D :
Une approche expérimentale.**

**Directeur de thèse :
Pascal HUGUET**

JURY

M. Ahmed CHANNOUF, Professeur, Université de Provence	Examineur
M. Jean-Claude CROIZET, Professeur, Université de Poitiers	Rapporteur
M. Pascal HUGUET, Directeur de recherche, CNRS & Université de Provence	Directeur
M. Olivier KLEIN, Professeur assistant, Université Libre de Bruxelles	Examineur
Mme Delphine MARTINOT, Professeur, Université Blaise Pascal – Clermont II	Rapporteur

REMERCIEMENTS

Je suis avant tout reconnaissant envers Pascal Huguet de m'avoir sollicité pour travailler sur un sujet aussi passionnant, moi qui ignorait jusque là tout de la psychologie sociale. Qu'il soit remercié pour la patience et la confiance dont j'ai toujours pu bénéficier.

Je tiens particulièrement à remercier Jean-Paul Caverni. Lorsque j'ai souhaité m'orienter vers la psychologie, il m'a accueilli au LPC avec toute la confiance et surtout la bienveillance dont il a toujours su témoigner. Je n'ai jamais eu l'occasion de douter de son soutien, de son écoute, et, surtout, de son humanité.

J'exprime toute ma reconnaissance envers Ahmmed Channouf, Jean-Claude Croizet, Olivier Klein, et Delphine Martinot qui me font l'honneur d'accepter d'évaluer ce travail de thèse.

Je remercie bien sûr l'ensemble des membres du LPC qui ont permis, à différents niveaux, l'aboutissement de ce travail. Plus particulièrement, merci à Michel Gonzalez et Isabelle Régner pour leurs encouragements et leurs conseils avisés. Merci à Aline Pélissier d'avoir remué ciel et terre pour me permettre d'obtenir les écrits de Sanders. Enfin mes remerciements les plus chaleureux à Catherine Thinus-Blanc pour sa relecture minutieuse et critique du manuscrit. Catherine tes interventions m'ont donné la motivation qui m'aurait manquée pour mener ce travail à bien.

Je remercie également Denis Cousineau pour ses encouragements et ses conseils directs qui m'ont permis de procéder sereinement aux analyses distributionnelles.

Mes pensées les plus chaleureuses pour les doctorants marseillais qui se sont avérés les meilleurs ami(e)s que l'on puisse espérer ! Bien sûr mes *roommates* Lënda et David dont le soutien a été inestimable, mais aussi Delphine, Caroline, Yousri, Lætitia, Annabelle, Magali, Emma, Stéphanie, Stéphanie, Pauline, Clément, François...

Je n'oublie pas non plus les doctorants nantais qui n'ont pas été en reste du point de vue de l'amitié : Linda « 12 », ainsi qu'Angélique, Aurélia, Sonia et Fabien. Je remercie également l'équipe du LabÉCD pour son accueil chaleureux, notamment Fabienne Colombel, Luc Jagot, Élise Renard, Yves Corson, Nadège Verrier, Françoise Michon, Françoise Piou et Lucas Bonnati.

Je remercie les étudiants qui m'ont apporté leur collaboration : Julien, Coralie, Aurélia et Jean-François. Une mention particulière pour Maude qui a déployé une énergie rare et une bonne humeur communicative. Maude, je te dois beaucoup. Je remercie également l'ensemble des étudiants à qui j'ai eu le plaisir d'enseigner, ils n'ont jamais su à quel point cette activité m'a rasséréiné.

Mes plus vifs remerciements aux collégien(ne)s et aux étudiant(e)s qui ont (librement) consenti à peiner sur les difficiles problèmes de rotation mentale que je leur soumettais. Qu'ils me pardonnent. Je remercie par la même occasion les collègues et les enseignants qui ont accepté de collaborer à l'étude 1.

Enfin, merci à tous ceux qui, hors du cercle scientifique, n'en ont pas moins apporté une contribution morale essentielle. Mes amis de toujours, Olivier, Daphnée, Julien, Stéphanie, Sébastien, Marie, Mathieu, Anne, Jean-Dominique, Christelle, et enfin, non des moindres, Xavier et Emmanuelle. Merci aussi à Sandra, Rosy, aux nantais qui m'ont fait paraître le séjour dans leur belle ville bien agréable, Guy, Jenn, Azia, Johnny, Mateusz et Pierre-Yves. Merci également à ceux qui prennent le « temps de voir », Marion, Nicole (merci pour la relecture), Arnaud, Odile, Pascal, Édith. Merci à ma mère, son énergie et son amour, à Benoît, à Clarisse. À la mémoire de mon père, enfin, à qui je dédie ce travail. Il reste ceux qui me pardonneront de ne pas les avoir nommés... dont un, cher à mon cœur, à l'origine de ce travail.

RÉSUMÉ

Les différences de sexe observées dans les tests de rotation mentale 3D sont connues pour être typiquement fortes et robustes, ancrées dans des mécanismes biologiques, et amplifiées au cours de la socialisation via l'intégration des rôles sexués. Ce consensus, cependant, ne prend pas en compte différents mécanismes liés à l'intervention de stéréotypes négatifs concernant les femmes dans le domaine visuo-spatial et susceptibles d'altérer leurs performances au cours même du test. Nous soutenons que ces mécanismes sont importants, sinon cruciaux, pour mieux comprendre l'origine des différences en question. Adossés notamment à la littérature sur la « menace du stéréotype », nos résultats expérimentaux (4 études, 1500 participants) nous conduisent à considérer la malléabilité de ces différences et leur ancrage dans des constructions culturelles (stéréotypes sociaux). Notre recherche contribue à ce titre à l'étude de la régulation sociale des processus cognitifs de base.

MOTS CLEFS

Activités visuo-spatiales, Différence de sexe, Régulation sociale, Rotation mentale, Stéréotypes sociaux

ABSTRACT

Social stereotypes and regulation of sex difference in 3D mental rotations : An experimental perspective.

The sex differences associated with 3D mental rotation tests are typically thought to be especially strong, robust, rooted in biological mechanisms, and amplified by gender role socialization. This consensus, however, does not take into account several mechanisms related to the intervention of negative stereotypes about women in the visuo-spatial domain, which may interfere with task performance. Here, we claim that these mechanisms are important otherwise crucial to understand better the origin of the differences in question. Based on the stereotype threat literature, our experimental findings (4 studies, 1500 participants) lead to consider the malleability of these differences and their anchoring in cultural knowledge (social stereotypes). As such, our research contributes to the study of the social regulation of basic operations of cognition.

KEYWORDS

Mental rotation, Sex difference, Social regulation, Social stereotype, Visuo-spatial activities

FORMATION DOCTORALE

Psychologie

Laboratoire de psychologie cognitive

CNRS UMR 6146

Université de Provence, Centre St-Charles

Pôle 3C — Comportement, Cerveau, Cognition — Bâtiment 9, Case D

3, Place Victor Hugo

13331 Marseille Cedex 1 France

Table des matières

Introduction générale	11
Chapitre 1 : Différences de sexe dans le domaine visuo-spatial	18
1.1 Introduction	18
1.2 Les différences de sexe dans le domaine visuo-spatial	19
1.2.1 Les aptitudes visuo-spatiale au cœur des différences de sexe ?	19
1.2.2 Les aptitudes verbales	21
1.2.3 Les différents tests visuo-spatiaux	23
1.3 La Rotation Mentale	26
1.3.1 Difficulté	27
1.3.2 Capacité en mémoire de travail	28
1.3.3 Stratégies	31
1.4 Facteurs biologiques et sociaux	37
1.4.1 Le substrat biologique des opérations de RM	38
1.4.2 Les facteurs sociaux	46
1.5 Conclusion	50
Chapitre 2 : Interférence du stéréotype de genre et rotation mentale	52
2.1 Introduction	52
2.1.1 Une alternative à l'approche macro-sociale	52
2.1.2 Les stéréotypes	53
2.2 Interférence, ou « menace », du stéréotype	54
2.2.1 Une Approche Situationnelle	54
2.2.2 Mécanismes	57
2.2.3 Différence de sexe visuo-spatiale et stéréotype	61
2.3 Stéréotype et rotation mentale	63
2.3.1 Sharps, Welton et Price (1993) : des précurseurs	63
2.3.2 Intervention du stéréotype en rotation mentale	67
2.4 Questions non résolues, nos objectifs	74
Chapitre 3 : Stéréotype et rotation mentale 3D, une approche expérimentale	77
3.1 Objectifs de l'étude 1	77
3.2 Méthode	82
3.2.1 Participants	82
3.2.2 Matériel	83

3.2.3	Procédure	85
3.2.4	Attentes	88
3.3	Résultats	90
3.3.1	Test de Rotation Mentale	90
3.3.2	Questionnaire	100
3.4	Discussion	109
3.5	Conclusion	115
Chapitre 4 : Interférence du stéréotype chez de jeunes adultes		116
4.1	Objectifs des études 2 et 3	116
4.2	Étude 2 : attentes stéréotypiques concernant le MRT	117
4.2.1	Méthode	119
4.2.2	Résultats	123
4.2.3	Discussion	126
4.3	Étude 3 : Intervention alternative à la falsification	129
4.3.1	Méthode	130
4.3.2	Résultats	132
4.3.3	Discussion	139
4.4	Conclusion	141
Chapitre 5 : Impact du stéréotype sur les mécanismes énergétiques . . .		143
5.1	Objectifs de l'étude 4	143
5.2	Modèle intégré du <i>stereotype threat</i>	144
5.2.1	Médiateurs du <i>stereotype threat</i>	144
5.2.2	Le modèle intégré de Schmader et coll. (2008)	146
5.3	Le modèle du stress de Sanders (1983)	148
5.3.1	Modèle énergétique et interférence du stéréotype	150
5.4	Méthode	155
5.4.1	Participants	155
5.4.2	Procédure	155
5.4.3	Attentes	158
5.5	Résultats	159
5.5.1	Taux de réponses correctes	159
5.5.2	Temps de réponse	161
5.5.3	Analyse distributionnelle	164
5.5.4	Rapport vitesse-précision	175
5.6	Discussion	177
5.7	Conclusion	180
Conclusion générale		181
Références		198
Index des auteurs		213
Index des sujets		219

Annexe A : Étude 1	222
Test de Rotation Mentale - Instructions	223
Test de Rotation Mentale - Section 1	225
Test de Rotation Mentale - Section 2	227
Questionnaire post-test - Instructions	229
Questionnaire post-test - Connaissance du stéréotype	230
Questionnaire post-test - Adhésion au stéréotype	232
Questionnaire post-test - Perception du MRT	234
Sexe-ratios - Sections 1 et 2 du MRT	235
Résultats sur l'échantillon global	236
Annexe B : Étude 2	237
Questionnaire sur les attentes concernant différents tests	238
Annexe C : Étude 3	251
Questionnaire post-test	252

Liste des figures

1.1	<i>Paper Folding Test</i> : le dessin représente une feuille de papier pliée puis percée. Il faut sélectionner la réponse qui correspond à la position correcte des trous une fois la feuille de papier dépliée. Ici, il s'agit du dessin A.	25
1.2	<i>Water Level Test</i> : une série de dessins représente des récipients inclinés, remplis de liquide. Il faut sélectionner le dessin qui représente correctement la surface du liquide (horizontale). Ici il s'agit du dessin n°1.	25
1.3	<i>Mental Rotation Test</i> : un modèle d'objet 3D est présenté à gauche, il faut sélectionner les deux représentations correctes du même objet — vues depuis différents angles — parmi les quatre options proposées. Ici il s'agit de la première et de la dernière proposition.	26
3.1	Sexe-ratios pour différentes tranches de performance. Chaque tranche représente 20% de l'effectif total.	92
3.2	Interaction <i>Sexe</i> \times <i>Instructions</i> sur l'échantillon total ($N=877$).	94
3.3	Interaction <i>Sexe</i> \times <i>Instructions</i> \times <i>Section</i> (score > 16).	98
4.1	Page tirée du questionnaire conçu pour mesurer les attentes des étudiants sur différents tests (ici IBT). Pour chaque problème donné en exemple, les participants indiquaient leurs attentes concernant la réussite de différentes populations cibles (ici seules quatre échelles sont données en exemple, voir l'annexe p. 238 pour la totalité des questions).	121
4.2	Interaction <i>Sexe</i> \times <i>Cible</i> (mesure des attentes au MRT)	124
4.3	Nombre de participantes (pourcentage) par niveau de performance au MRT.134	
4.4	Interaction <i>Instructions</i> \times <i>Section</i> sur l'échantillon total.	136

5.1	Modèle intégré de Schmader et coll. (2008) montrant les différents mécanismes impliqués dans la baisse de performance due au <i>stereotype threat</i> . . .	147
5.2	Modèle du stress et de la performance de Sanders (1983) montrant l'interaction entre les différentes étapes du traitement cognitif de l'information, et les processus énergétiques modulant l'efficacité de ce traitement sous la supervision d'un processus d'évaluation.	149
5.3	Déroulement d'un essai (MRT informatisé).	157
5.4	Affichage de l'écran censé activer l'identité de sexe des participantes. Il était demandé aux sujets de sélectionner (avec la souris) le symbole correspondant à leur sexe.	158
5.5	Précision (pourcentage réponse correcte) - Interaction <i>Cursus</i> \times <i>Angles</i> . .	160
5.6	Temps de Réponse - Effet de l'amplitude de l'angle.	162
5.7	Temps de Réponse - Interaction <i>Instruction</i> \times <i>Cursus</i> \times <i>Angle</i>	163
5.8	La convolution d'une distribution normale avec les paramètres $\mu = 500$ et $\sigma = 100$ (A) avec une distribution exponentielle avec le paramètre $\tau = 250$ (B) résulte en une distribution ex-Gaussienne (C) (exemple tiré de Lacouture et Cousineau, 2008).	166
5.9	Paramètre τ , interactions <i>Cursus</i> \times <i>Instructions</i> pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de τ dans chaque modalité.	169
5.10	Paramètre μ , interactions <i>Cursus</i> \times <i>Instructions</i> pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de μ dans chaque modalité.	171
5.11	Paramètre σ , interactions <i>Cursus</i> \times <i>Instructions</i> pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de σ dans chaque modalité.	174
5.12	Pourcentages cumulés des réponses correctes en fonction de la rapidité des réponses.	176
A.1	Sexe-ratios pour différentes tranches de performance. Chaque tranche représente 20% de l'effectif total.. . . .	235
A.2	Interaction <i>Sexe</i> \times <i>Instructions</i> \times <i>Section</i> (échantillon total).	236

Liste des tableaux

3.1	<i>Taille de effet l'effet de sexe (d de Cohen) pour chaque section du MRT et pour l'ensemble du test, par type d'instruction (échantillon total, N= 877).</i>	95
3.2	<i>Taille de l'effet de sexe (d de Cohen) pour chaque section du MRT et pour l'ensemble du test, par type d'instruction (échantillon limité aux meilleurs scores, N = 304).</i>	100
4.1	<i>Attentes concernant la réussite des hommes et des femmes à différents tests.</i>	126
4.2	<i>Performance des étudiantes aux deux parties du MRT en fonction des instructions.</i>	135
5.1	<i>Analyses (ANOVA) pour le paramètre τ, estimé pour chaque type de paire d'objets (Identiques vs. Différents). Les effets statistiquement significatifs sont en gras.</i>	168
5.2	<i>Analyses (ANOVA) pour le paramètre μ, estimé pour chaque type de paire d'objets (Identiques vs. Différents). Les effets statistiquement significatifs sont en gras.</i>	170
5.3	<i>Analyses (ANOVA) pour le paramètre σ, estimé pour chaque type de paire d'objets (Identiques vs. Différents). Les effets statistiquement significatifs sont en gras.</i>	173

Introduction générale

L'objectif de cette thèse est d'apporter un éclairage psychosocial sur les mécanismes susceptibles d'expliquer les différences de sexe habituellement observées sur les tests visuo-spatiaux. Nous nous focaliserons sur les tests de rotation mentale, associés dans la littérature en psychologie cognitive à l'une des différences de sexe (en faveur des hommes) les plus robustes (Halpern, 2000 ; Linn & Petersen, 1985 ; Voyer, Voyer, & Bryden, 1995). Notre thèse est que cette différence reflète, au moins en partie, l'intervention d'un stéréotype négatif à l'égard des femmes dans le domaine visuo-spatial. Cette intervention fera donc l'objet de nos travaux expérimentaux dans la lignée de la littérature impulsée par Steele et ses collaborateurs sur les stéréotypes interférents (Steele & Aronson, 1995 ; Steele, 1997).

Nos travaux se situent ainsi à l'interface de deux champs de recherche en psychologie, qui n'ont été mis en relation que récemment (Hausmann, Schoofs, Rosenthal, & Jordan, 2009 ; Martens, Johns, Greenberg, & Schimel, 2006 ; M. S. McGlone & Aronson, 2006 ; Moè & Pazzaglia, 2006 ; Wraga, Duncan, Jacobs, Helt, & Church, 2006 ; Wraga, Helt, Jacobs, & Sullivan, 2007). Les résultats de cette littérature, bien qu'encourageants, ne sont néanmoins pas suffisants pour conclure à l'intervention du stéréotype dont nous faisons l'hypothèse. Le but de nos propres travaux est donc de poursuivre cette réflexion et de préciser l'impact réel de ce stéréotype sur la performance et les processus impliqués dans le domaine des rotations mentales.

Si les différences de sexe dans le domaine visuo-spatial sont bien admises (Halpern, 2000 ; Kimura, 1999 ; Maccoby, 1966 ; Maccoby & Jacklin, 1974), leur amplitude en revanche fait encore débat (Halpern et coll., 2007). La taille des différences en question varie en effet selon les tests utilisés : les différences les plus importantes (toujours en faveur des hommes) concernent les tests de rotation mentale d'objets en trois dimensions (cf. chapitre 1 p. 26, Linn & Petersen, 1985 ; Voyer et coll., 1995). Dans les faits, la différence en rotation mentale 3D est de si grande taille (de l'ordre d'un écart-type), et elle est si robuste (à travers les cultures et les époques, cf. Peters et coll., 1995 ; Voyer et coll., 1995), que certains auteurs en ont conclu qu'il s'agissait là de la preuve d'une supériorité masculine biologiquement inscrite (Geary, 1995 ; Kimura, 1999). Cette position est confortée par des résultats montrant chez les femmes une corrélation positive entre le niveau de testostérone et la réussite aux tests de rotation mentale 3D, et une corrélation négative avec le niveau d'œstrogène dans le sang au moment du test (Hausmann, Slabbekoorn, Van Goozen, Cohen-Kettenis, & Güntürkün, 2000). En outre, des travaux fondés sur la technique d'imagerie par résonance magnétique fonctionnelle (IRMf) ont montré que les individus des deux sexes ne mobilisaient pas nécessairement les mêmes zones cérébrales au cours d'une tâche de rotation mentale (Jordan, Wüstenberg, Heinze, Peters, & Jäncke, 2002 ; Weiss et coll., 2003).

Une des régions cérébrales impliquées dans cette différence de sexe est le cortex pré-frontal, en particulier orbitofrontal. Or, cette région est connue pour son implication au niveau des processus de prise de décision, notamment en rapport à soi, aux émotions (risque) et à la mémoire de travail (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009 ; Kringelbach, 2005 ; LoPresti et coll., 2008). Il n'est donc pas exclu que des différences d'activation cérébrale réfrentent au moins autant des effets de contexte que des différences neuro-biologiques fondamentales entre hommes et femmes dans le traitement de l'information visuo-spatiale. En effet Wraga et coll. (2007) montrent que la seule manipulation de croyances sur l'efficacité relative des deux sexes en matière de rotation mentale suffit à produire des différences d'activation cérébrale, s'agissant en particulier des régions

associées aux émotions (région rostro-ventrale gauche du cortex cingulaire antérieur, gyrus orbital droit). Il semble donc que les différences de sexe en matière de rotation mentale ne soient pas indépendantes des attentes ou croyances forgées par les sujets sous l’emprise de leur appartenance de sexe et des stéréotypes afférents.

Dans le champ de la psychologie sociale, Steele (1997) a théorisé et identifié, sous le nom de *stereotype threat*, les conséquences comportementales que les réputations d’infériorité sont susceptibles de générer auprès des individus qui en sont la cible. Ces conséquences s’expriment notamment par la production de performances en deçà des compétences réelles des individus, d’où des inégalités de performance entre groupes sociaux assez souvent conformes au stéréotype lui-même. Précisons d’ores et déjà que ce phénomène n’implique pas nécessairement un mécanisme d’intériorisation du stéréotype au cours de la socialisation. En fait, la grande originalité de l’hypothèse avancée par Steele et ses collaborateurs est d’envisager une perturbation transitoire du fonctionnement cognitif sous l’effet de la récupération en mémoire de connaissances liées au stéréotype en situation d’évaluation. Lorsque le test est supposé diagnostique des compétences en jeu dans le stéréotype (p. ex. compétences verbales, compétences en mathématiques), les individus — en particulier ceux paradoxalement les plus compétents —, pourraient alors craindre de confirmer la mauvaise réputation de leur groupe d’appartenance. Or, en consommant des ressources cognitives au détriment du test lui-même, cette crainte est précisément susceptible de conduire à un échec en effet de nature à confirmer le stéréotype négatif. Dans l’approche de Steele (1997), la simple connaissance du stéréotype constitue ainsi une condition suffisante à la production d’une réalité qui au final lui est conforme.

La mise en évidence du phénomène invoqué par Steele et ses collaborateurs est relativement récente — l’étude *princeps* de Steele et Aronson date de 1995. Elle a cependant suscité au cours des 15 dernières années un nombre très important de publications, qui pour la plupart ont attesté de la robustesse du phénomène en question. Ce phénomène semble très général, « ubiquitaire », car il concerne les membres de différents groupes so-

ciaux (femmes, hommes, afro-américains, latino-américains, asiatiques, personnes âgées, élèves d'origine modeste, etc.) et touche à tous les domaines spécifiés par le stéréotype ciblant le groupe social en question (infériorité des femmes en mathématiques, des afro-américains en matière d'intelligence, des personnes âgées dans le domaine de la mémoire, etc.).

Néanmoins les études qui ont cherché, au sein du paradigme du *stereotype threat*, à mettre en évidence le rôle des stéréotypes dans l'expression de la différence de sexe en rotation mentale n'ont pas réellement permis d'aboutir. En effet soit ces études montraient seulement une altération de la performance des *hommes* (Wraga et coll., 2006), ce qui ne permet pas de conclure que le stéréotype empêche les *femmes* de réussir, soit les effets rapportés étaient trop faibles pour annuler la différence de sexe (Hausmann et coll., 2009 ; M. S. McGlone & Aronson, 2006). Par ailleurs, certaines études (p. ex. Moè & Pazzaglia, 2006 ; Wraga et coll., 2007) ne permettaient simplement pas de comparer directement les performances des deux sexes, et donc de mesurer l'influence du stéréotype sur la différence de sexe si souvent observée par ailleurs. Enfin, du fait de leur non-respect de la méthodologie en vigueur dans le paradigme du *stereotype threat*, d'autres travaux encore (Martens et coll., 2006 ; Moè & Pazzaglia, 2006 ; Wraga et coll., 2006, 2007) n'ont pas davantage permis d'établir le rôle causal des stéréotypes dans le domaine des différences de sexe en matière de rotations mentales.

En dépit des difficultés et autres lacunes de nature essentiellement méthodologiques brièvement évoquées ci-dessus, la thèse défendue dans nos travaux est que le phénomène invoqué par Steele dans les années 90 est bien en grande partie à l'origine même des différences de sexe observées dans le domaine spécifique des rotations mentales. Nos travaux visent donc à réexaminer l'idée d'une causalité ancrée dans l'intervention d'un stéréotype défavorable aux femmes dans le domaine visuo-spatial.

Cinq chapitres structurent ce rapport de thèse. Nous consacrerons le *premier chapitre* à une revue de questions concernant l'étude des différences de sexe au niveau cognitif.

Nous montrerons dans un premier temps combien les outils de mesure de la différence de sexe se sont affinés et ont permis de passer d'une vision relativement monolithique de ces différences à une approche beaucoup plus fine. Ainsi, différents facteurs ont été mis à jour (type de tâches, type de processus mobilisés) pour mieux caractériser ces différences. Ceci nous permettra d'explicitier les raisons pour lesquelles les tests de rotation mentale occupent une place aussi singulière dans l'étude des différences de sexe. Nous tenterons au passage une synthèse des résultats concernant les deux grandes catégories de facteurs explicatifs des différences observées, en montrant que les facteurs biologiques ne peuvent à eux seuls expliquer tous les aspects des différences observées, et les obstacles — notamment méthodologiques — rencontrés dans la recherche des déterminants psychosociaux. En effet, ces derniers sont majoritairement abordés sous l'angle de l'hypothèse de la socialisation (une influence diffuse de l'environnement social sur le long terme). Ils sont donc généralement appréhendés sous la forme de facteurs invoqués dans des études simplement corrélationnelles, ce qui ne permet pas de conclure quant à leur rôle causal.

Une seconde synthèse, concernant cette fois la théorie du *stereotype threat*, fera l'objet du *chapitre 2*. Après un exposé détaillé de la théorie de Steele, nous nous arrêterons sur une procédure dite de *falsification explicite du stéréotype*, testée par Spencer, Steele, et Quinn dès 1999, et, nous le verrons, centrale dans nos propres travaux. Nous procéderons ensuite à une revue critique des différentes études déjà conduites à propos de l'implication des stéréotypes de genre dans les différences hommes / femmes en matière de rotation mentale. Selon nous, les études antérieures dans ce domaine n'ont pas été très concluantes en raison, notamment, de la négligence d'un point pourtant central dans la théorie de Steele. En effet, selon cette théorie, les individus les plus vulnérables à l'interférence du stéréotype sont ceux pour qui la réussite à la tâche présente un enjeu, en particulier ceux les plus compétents dans le domaine évalué et qui par conséquent ont beaucoup à perdre en cas d'échec.

Partant de ce principe, nous présenterons dans le *chapitre 3* une étude conçue pour mesurer l'effet de la *falsification* du stéréotype à différents niveaux de compétence dans l'activité cible. Cette étude, menée sur un échantillon de très grande taille ($N = 877$), montre que la falsification du stéréotype en effet annule la différence de sexe chez les individus les plus compétents. Cette étude montre également qu'il existe une dynamique temporelle du phénomène : nous constaterons alors que, dans les conditions où le stéréotype est supposé activé, les différences apparaissent essentiellement dans la seconde section du test. Cette dynamique est compatible avec l'idée selon laquelle la confrontation progressive à la difficulté du test joue un rôle capital dans l'intervention du stéréotype, en poussant les sujets à douter de leur réussite (incertitude anxiogène) ou à nourrir des pensées négatives interférentes (Cadinu, Maass, Rosabianca, & Kiesner, 2005).

Une deuxième étude, présentée au cours du *chapitre 4*, nous permettra de vérifier que des individus des deux sexes confrontés à un test de rotation mentale nourrissent en effet spontanément des attentes compatibles avec le stéréotype quant au succès relatif des hommes et des femmes sur ce test. Car si l'on accepte généralement dans la littérature l'existence d'un stéréotype concernant les aptitudes visuo-spatiales des femmes, très peu d'études ont jusqu'ici mesuré directement l'existence de telles attentes face à des tests visuo-spatiaux (Blakemore, Berenbaum, & Liben, 2009 ; Deaux, 1985 ; Halpern, 2000 ; Robert, 1990).

Une troisième étude visera surtout à évaluer l'efficacité d'une induction alternative à celle dite de la falsification et consistant à installer les sujets (au moment du test) dans une comparaison sociale non-menaçante, en l'occurrence avec des personnes âgées. Nous verrons qu'en effet l'influence de cette comparaison est bel et bien perceptible au niveau de la dynamique temporelle évoquée antérieurement.

Enfin, notre quatrième et dernière étude, objet du *chapitre 5*, testera les processus sous-jacents à l'interférence du stéréotype grâce au recueil de temps de réponse (TR) habituellement ignorés dans les travaux adossés à l'hypothèse de Steele (1997). Nous

comparerons en outre les performances d'étudiantes en sciences et en lettres dans une condition de falsification et une condition contrôle, avec l'hypothèse d'un effet positif de la falsification limité aux femmes issues de filières scientifiques, *a priori* plus identifiées au domaine visuo-spatial. Mais surtout, une analyse originale, focalisée sur la forme des distributions des TR permettra d'inférer que le stéréotype interfère avec des processus liés à l'effort et au contrôle des ressources attentionnelles plutôt qu'avec le traitement de l'information visuo-spatiale *per se*. Ces nouveaux résultats confirmeront des prédictions élaborées à partir du modèle du stress et de la performance de Sanders (1983), compatible avec les hypothèses actuelles au sujet des mécanismes responsables du *stereotype threat* (Schmader, Johns, & Forbes, 2008). Ce modèle stipule que l'intervention de facteurs liés au contrôle attentionnel se traduit par un allongement des distributions des TR (c.-à-d. un allongement des temps les plus longs). Nos résultats, nous le verrons, sont assez conformes à cette hypothèse. L'approche esquissée dans cet ultime chapitre permet selon nous non seulement d'éclairer la nature de l'interférence cognitive liée à l'intervention du stéréotype au cœur de nos travaux, mais aussi de faire le lien avec la littérature concernant la théorie de l'effort et de la gestion des ressources cognitives.

Nous concluons par une synthèse de nos résultats et discuterons à la fois leur apport spécifique dans le champ des études sur les différences de sexe et leurs enjeux sociétaux. Des perspectives de recherche seront également suggérées.

Chapitre 1

Différences de sexe dans le domaine visuo-spatial

1.1 Introduction

L'objectif de ce chapitre est d'offrir une vision d'ensemble des connaissances actuelles sur les différences de sexe au niveau du traitement de l'information visuo-spatiale. Une revue de la littérature de ces trente dernières années justifiera que nous nous focalisions sur la rotation mentale (RM), les tests de RM montrant les différences de sexe les plus importantes et donc les plus informatives. Nous ferons ensuite le point sur les processus cognitifs (stratégies, mobilisation de la mémoire de travail, etc.) qui jouent un rôle en RM. Enfin, sur cette base, nous aborderons les principales hypothèses explicatives de cette différence de sexe en RM. Ces hypothèses mettent l'accent sur des facteurs d'origine biologique (dimorphisme neuro-anatomique, latéralisation fonctionnelle, etc.) ou sociale (socialisation, éducation, stéréotypes, etc.), ou encore sur l'interaction de ces deux facteurs sur le long terme (modèles bio-psycho-sociaux). Cette revue est nécessaire afin d'apprécier pleinement l'originalité de l'approche que nous proposons, qui met l'accent sur le contexte

social et la situation de test, et que nous présenterons dans le second chapitre de cette thèse.

1.2 Les différences de sexe dans le domaine visuo-spatial

Maccoby et Jacklin (1974) ont réalisé l'une des premières synthèses de la littérature sur les différences de sexes, portant sur près de 1600 études dans différents domaines de la psychologie (cognition, développement, socialisation, personnalité, etc.). Ces auteurs ont conclu que les différences n'étaient pratiquement décelées qu'au niveau cognitif¹ : les femmes montreraient de meilleures capacités verbales que les hommes tandis que ces derniers excellerait dans les tâches se rapportant aux mathématiques et au traitement de l'information visuo-spatiale. Ces résultats ont depuis structuré une large part de l'approche cognitive des différences entre hommes et femmes, que l'on oppose généralement sur la base de leurs capacités verbales, visuo-spatiales et en mathématiques (Halpern, 2000 ; Kimura, 1999).

1.2.1 Les aptitudes visuo-spatiale au cœur des différences de sexe ?

Si les différences de sexe semblent s'exprimer essentiellement dans trois domaines — verbal, visuo-spatial et mathématiques — ces aptitudes ne sont pas totalement indépendantes les unes des autres. Par exemple, les hommes surpassent les femmes dans divers sous-domaines des mathématiques (raisonnement, résolution de problème et géométrie, cf. Halpern et coll., 2007) qui semblent mobiliser et intégrer des processus visuo-spatiaux

1. Maccoby et Jacklin (1974) n'ont retenu qu'une seule nette différence dans un autre domaine, qui concerne un trait de personnalité : l'agressivité.

plus élémentaires (visualisation, transformation, rotation spatiales, etc.) : la supériorité dans le domaine visuo-spatial peut donc, de fait, conduire à une supériorité dans le domaine mathématique. Plusieurs revues de questions et méta-analyses concluent à l'existence de corrélations élevées entre réussite dans le domaine mathématique et les capacités visuo-spatiales (Lubinski & Benbow, 2006 ; Webb, Lubinski, & Benbow, 2007).

Casey, Nuttall, Pezaris, et Benbow (1995) ont testé plus spécifiquement l'hypothèse d'une médiation de la différence de sexe en mathématiques par la différence dans les aptitudes visuo-spatiales. Ces auteurs ont proposé conjointement un test spatial de rotation mentale (RM) et un test standard de mathématiques (SAT-M) et à des groupes d'étudiants des deux sexes et de différents niveaux académiques. Des corrélations significatives entre performance en RM et en mathématiques (entre $r = .35$ et $r = .38$) ont été observées dans tous les groupes de femmes. Parallèlement, les hommes ont montré leur supériorité en mathématiques et en RM, sauf chez les étudiants de plus faible niveau. Pour établir le sens de la causalité entre les performances en RM et en mathématiques, les auteurs ont procédé à des ANCOVA, testant l'effet de sexe en RM en ajustant les scores obtenus au SAT-M, et vice-versa. Ces analyses ont montré que la différence de sexe en mathématiques disparaissait lorsque l'effet de sexe en RM était contrôlé dans les groupes qui présentaient de bons résultats scolaires. À l'inverse, l'ajustement des scores au SAT-M n'avait aucune incidence sur l'effet de sexe en RM. Un lien de nature causale semble donc exister entre les performances en RM et en mathématiques, une bonne aptitude en RM expliquant la réussite des sujets les plus doués en mathématiques, notamment féminins. Il apparaît donc que les capacités de traitement de l'information visuo-spatiales occupent une place déterminante dans l'étude des différences de sexe, car elles sont susceptibles d'expliquer ces mêmes différences dans d'autres domaines, plus intégrés, comme les mathématiques. En outre, cette médiation n'apparaît, comme les différences de sexe en question, que chez les sujets d'un bon niveau académique. Nous reviendrons sur ce dernier point, remarquable dans ce genre d'études : les différences sont observées essentiellement au niveau

des échantillons les plus sélectifs, c'est-à-dire pour les niveaux de performances les plus élevés (Hedges & Nowell, 1995 ; Halpern, 2000 ; Halpern et coll., 2007).

En dernière analyse, il apparaît qu'une meilleure compréhension de la différence de sexe dans le domaine visuo-spatial éclairerait également la différence observée dans d'autres domaines cognitifs, comme les mathématiques. Le traitement de l'information visuo-spatiale est d'assez bas niveau, et semble donc jouer un rôle fondamental dans bon nombre d'activités cognitives plus intégrées (mathématiques, raisonnement, navigation, etc.). Expliquer les différences de sexe survenant à ce niveau est donc essentiel, d'autant que de nombreux indices indiquent que les processus à l'origine de cette différence sont relativement complexes. Tous les hommes et toutes les femmes ne semblent pas montrer de ce point de vue des comportements strictement identiques. Dans les sections suivantes, nous détaillons et définissons plus précisément ce domaine visuo-spatial, qui est loin d'être monolithique.

1.2.2 Les aptitudes verbales

Les aptitudes visuo-spatiales sont définies comme « *la capacité de se représenter, de transformer, générer et rappeler une information symbolique non-verbale*² » (Linn & Petersen, 1985). Ces aptitudes visuo-spatiales sont donc appréhendées en opposition avec les aptitudes verbales. Or cette opposition se retrouve au niveau des différences de sexe : non seulement il existe une idée reçue selon laquelle le langage est le domaine féminin par excellence tandis que le domaine spatial est celui des hommes, mais cette idée semble corroborée scientifiquement (Halpern, 2000). Avant de présenter en détail les résultats qui concernent les aptitudes visuo-spatiales, nous allons rapidement évoquer quelques résultats concernant les aptitudes verbales.

Selon Maccoby et Jacklin (1974), McCarthy (1954) concluait déjà dans un premier travail de synthèse à un développement plus précoce des aptitudes verbales des filles.

2. Bien sûr, il est sous-entendu que l'information en question est d'origine visuelle.

Maccoby et Jacklin (1974), après Maccoby (1966), concluent que les différences de sexes deviennent substantielles passé l'âge de 11 ans. Cependant, si une différence peut être statistiquement significative, il est important également de déterminer si l'écart en question est de grande taille ou non. Or ce n'est qu'en 1981 que Hyde envisagea de calculer la part de la variance totale expliquée par le facteur sexe, notée ω^2 ³ et d'autre part le calcul de la différence standardisée entre les moyennes, c'est-à-dire le d de J. Cohen (1977)⁴. Pour les tests verbaux montrant une différence de sexe significative, Hyde conclut que $\omega^2 = .01$ et $d = -.24$, ce qui correspond en dernière analyse à un effet de sexe de petite taille. Finalement, Hyde et Linn (1988) ont montré sur 165 études que les femmes affichaient un avantage modéré dans les tâches de production verbale ($d = -.33$), et un avantage petit ou nul dans les autres tests (anagrammes, compréhension, rédaction, etc.), voire même favorisant les hommes pour les analogies ($d = .16$). Ainsi c'est finalement un type très particulier de tâches verbales qui permet de reproduire des différences à l'avantage des femmes, encore que cet avantage soit relativement faible.

Il apparait donc que les tests verbaux sont finalement assez peu discriminants, voire pas du tout, comparés aux test visuo-spatiaux où les différences de sexe sont plus nettes. L'étude des aptitudes visuo-spatiales a ainsi permis de développer un regard plus fin sur la différence de sexe, et se révèle alors un champ d'étude privilégié pour comprendre l'origine de ces différences.

3. ω^2 est dans ce cas la proportion de variance total expliquée par le sexe, calculé par la formule $\omega^2 = \frac{\text{SommedesCarrés}_{\text{sexe}} - \text{CarréMoyen}_{\text{erreur}}}{\text{CarréMoyen}_{\text{erreur}} + \text{SommedesCarrés}_{\text{total}}}$

4. d est un indice statistique introduit par J. Cohen (1977) pour estimer simplement la taille d'un effet observé entre deux groupes. La formule pour calculer la différence entre les deux sexes est : $d = \frac{M_H - M_F}{\sigma_{\text{total}}}$. Il s'agit simplement de standardiser la différence entre les deux groupes, en calculant le rapport de la différence entre les moyennes des deux échantillons comparés sur l'écart-type de la population totale (estimé à partir de l'écart type calculé après réunion des deux groupes). $d = 1$ signifie p. ex. que la distance entre les deux moyennes est égale à un écart-type. J. Cohen (1992) propose de qualifier d de *grand* lorsque $d \approx .80$, de *modéré* lorsque $d \approx .50$ et enfin de *petit* lorsque $d \approx .20$.

1.2.3 Les différents tests visuo-spatiaux

Les études dans le domaine visuo-spatial confirment également l'idée reçue que les femmes obtiennent des résultats inférieurs à ceux des hommes. Maccoby et Jacklin (1974) recensent 78 études qui montrent une infériorité féminine à une grande variété de tests spatiaux : labyrinthes, estimation de l'orientation d'un trait, extraction de figures simples cachées dans des dessins complexes, ou encore les problèmes spatiaux issus de tests standardisés (*Differential Aptitude Test*, *Wechsler Intelligence Scale for Children*). Ces tests, très différents, semblent mobiliser des opérations cognitives assez variées. Mesurent-ils tous l'efficacité des mêmes processus ?

Afin de mieux caractériser les différences observées, Hyde (1981) a proposé, comme pour les tests verbaux, de calculer la taille des effets de sexe. Elle a analysé séparément les tests qui demandaient une décomposition « analytique » de l'information visuo-spatiale (par exemple l'extraction d'une figure cachée) et les tests qui ne nécessitaient pas un tel traitement (par exemple la comparaison de l'orientation de deux lignes). Pour ces derniers, la part de variance expliquée par le sexe était de $\omega^2 = .043$ avec une taille d'effet modérée ($d = .45$), et pour les tests avec une composante analytique, $\omega^2 = .025$ et $d = .51$. Nous pouvons donc retenir que premièrement, les valeurs de d calculées pour les deux types de tests visuo-spatiaux étaient pratiquement identiques. Il n'était donc pas spécialement pertinent de les distinguer selon cette soi-disant composante « analytique ». Deuxièmement, ces valeurs de d indiquaient néanmoins que l'avantage des hommes sur les femmes dans les tests visuo-spatiaux était deux fois plus important que l'avantage des femmes sur les hommes dans les tests verbaux. La plus grande différence de sexe apparaissait donc dans le domaine visuospatial, même si finalement le sexe n'expliquait au mieux que 5% de la variance totale de la performance à ces tests.

À la suite de Maccoby et Jacklin (1974) et Hyde (1981), des calculs plus systématiques des tailles d'effet ont permis à Linn et Petersen (1985) de catégoriser plus finement les

tests visuo-spatiaux. Linn et Petersen ont observé que ces tests étaient très variés et ne mobilisaient pas les mêmes opérations visuo-spatiales. Or la différence de sexe pouvait être due au fait que les hommes réalisaient certaines de ces opérations mieux que les femmes. Il était donc possible que pour certaines opérations visuo-spatiales l'écart entre les sexes fût important, mais plus réduit pour d'autres. Linn et Petersen ont testé cette hypothèse en dégageant de l'analyse de 172 études trois grands ensembles de tâches visuo-spatiales, qui correspondaient à des opérations visuo-spatiales particulières, associés à une taille d'effet spécifique.

Tests de visualisation spatiale. Cet ensemble regroupe les tâches qui mobilisent les capacités de *visualisation spatiale*, c'est-à-dire une capacité à manipuler des objets dans l'espace et leurs relations entre eux (cf. le *Paper Folding Test*, fig. 1.1 p. 25). Les effets observés dans cette catégorie sont typiquement *petits* ($d = .13$). Du point de vue cognitif, l'information spatiale est manipulée dans ces tests en plusieurs étapes successives, d'où la possibilité de réaliser des choix stratégiques pour atteindre la solution. Cette composante stratégique peut expliquer la petite taille de l'effet : ce n'est pas tant les capacités visuo-spatiales qui sont critiques pour résoudre ces problèmes que les capacités d'adaptabilité ou « fluidité » (cf. théorie de l'intelligence, p. ex. Cattell, 1971). Or le sexe n'est pas un facteur critique pour cette capacité à s'adapter. D'autres tests qui appartiennent à cette catégories sont : le *Embedded Figure Test* (extraire une forme simple d'un dessin complexe), le *Paper Form Board* (imaginer une forme géométrique 3D à partir de la présentation d'un patron en 2D), le *Block Design Test* (reproduire un dessin en assemblant des cubes colorés), etc.

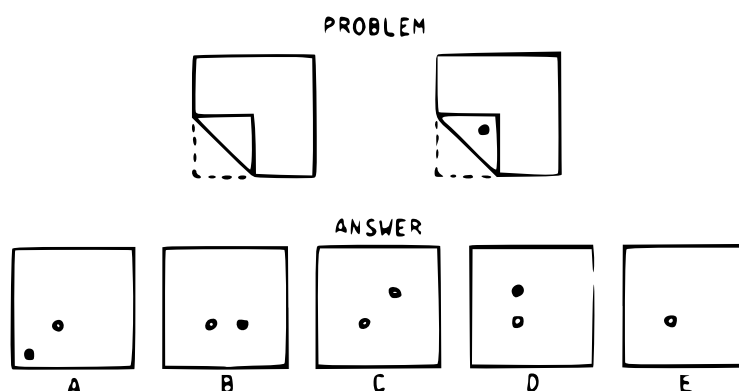


FIGURE 1.1 — *Paper Folding Test* : le dessin représente une feuille de papier pliée puis percée. Il faut sélectionner la réponse qui correspond à la position correcte des trous une fois la feuille de papier dépliée. Ici, il s'agit du dessin A.

Tests de Perception Spatiale. La seconde catégorie que ces auteurs ont examinée regroupe les tests qui demandent d'établir des relations spatiales entre différents objets, et ce malgré la présence de distracteurs (cf. le *Water Level Test*, fig. 1.2 p. 25). Dans le cas du WLT, le fait que le récipient soit incliné livre des informations interférentes (lignes en biais, etc.) qui peuvent gêner la perception de l'horizontale, nécessaire pour répondre correctement. Linn et Petersen appellent cette famille *tâches de perception spatiale*. La taille d'effet observée pour ce genre de tâche est modérée, ($d = 0.44$). La résolution de ces problèmes fait moins appel à la capacité d'élaborer une stratégie efficace qu'à celle d'inhiber une information distractive, au profit de l'information issue de la perception de l'orientation de son propre corps (proprioception). Un autre test de perception spatiale courant est le *Rod and Frame Test*, où il faut orienter une baguette à la verticale alors que l'on est placé devant un cadre rectangulaire incliné.

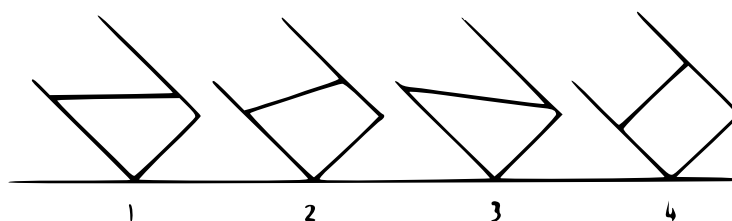


FIGURE 1.2 — *Water Level Test* : une série de dessins représente des récipients inclinés, remplis de liquide. Il faut sélectionner le dessin qui représente correctement la surface du liquide (horizontale). Ici il s'agit du dessin n°1.

Tests de Rotation Mentale. Enfin, la dernière catégorie regroupe les tâches qui nécessitent d’imaginer des rotations d’objets (en trois dimensions) ou de figures (en deux dimensions). La figure 1.3 (p. 26) présente un exemple de problème de rotation en trois dimensions issu du *Mental Rotation Test* (MRT) de Vandenberg et Kuse (1978)⁵. Les tests de ce type sont sensés mesurer la capacité de *rotation mentale* (RM). La RM est souvent présentée comme un processus unitaire et fondamental de traitement de l’information spatiale. Cette aptitude nous est absolument nécessaire pour reconnaître un objet ou une scène depuis des angles de vue différents, ce qui arrive notamment lorsque nous sommes mobiles dans un environnement. Linn et Petersen (1985) ont montré que c’est pour ce type de test que l’on observe les effets de sexe les plus importants, avec en moyenne $d = 0.73$.

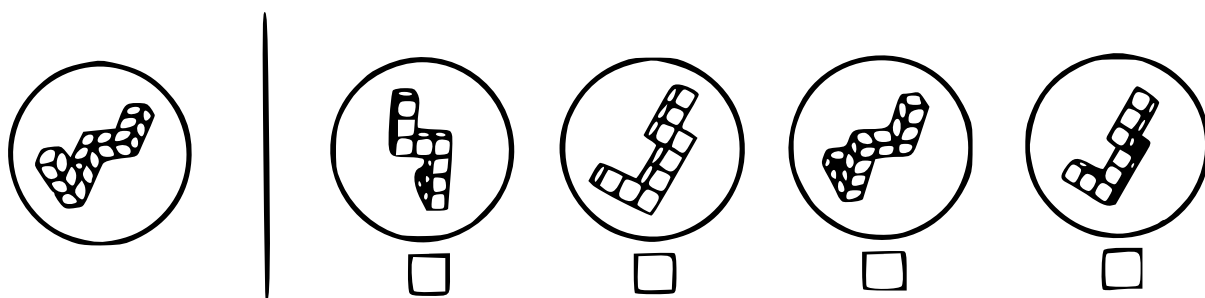


FIGURE 1.3 — *Mental Rotation Test* : un modèle d’objet 3D est présenté à gauche, il faut sélectionner les deux représentations correctes du même objet — vues depuis différents angles — parmi les quatre options proposées. Ici il s’agit de la première et de la dernière proposition.

1.3 La Rotation Mentale

Les tests de RM sont donc ceux qui montrent le plus nettement l’infériorité des femmes dans le domaine visuo-spatial. Il est question de « rotation mentale » car il est largement admis que pour résoudre ce type de problème la majorité des sujets procède à la rotation d’une image mentale de l’objet, d’où l’analogie avec la rotation d’un objet matériel. Non seulement cette idée est confirmée par les sujets qui déclarent faire « tourner » mentale-

5. Le test complet est disponible en annexe, p. 223

ment une représentation imagée de l'objet, mais aussi par des indices objectifs, comme les temps de réponse. Ainsi, Shepard et Metzler (1971) ont observé que des sujets qui devaient comparer deux représentations d'objets en 3D (du même type que ceux présentés fig. 1.3⁶) vus sous des angles différents, répondaient d'autant plus lentement que la comparaison nécessitait une rotation de grande amplitude. De là, ces auteurs ont pu déterminer l'existence d'une relation de proportionnalité entre la durée des temps de réponse et la magnitude de l'angle de rotation, preuve que l'opération mentale en jeu était analogue à une rotation physique. La rotation mentale est donc caractérisée par une certaine vitesse de traitement, les calculs indiquant qu'en moyenne il faut une seconde pour réaliser une rotation de 60°.

1.3.1 Difficulté

La RM est donc un processus cognitif bien identifié et singulier dans le domaine des différences de sexe car très peu de tests cognitifs permettent d'observer une taille d'effet aussi importante. Il s'agit donc de tâches de choix pour étudier les différences de sexe, particulièrement les tests de rotation en trois dimensions. En effet Linn et Petersen (1985) constatent que les écarts observés lors d'études basées sur le MRT de Vandenberg et Kuse (1978) sont homogènes et très importants ($d = .93$) à tous les âges, tandis que les tailles d'effets observés sur des tests en deux dimensions (par exemple le *Flags and Cards* tiré de French, Ekstrom, & Price, 1963, ou la sous-partie spatiale du *Primary Mental Ability*, Thurstone & Thurstone, 1941) ne permettaient pas toujours d'observer une différence de sexe, d'autant que l'effet, quand il existe, est seulement modéré ($d = .26$). Collins et Kimura (1997) ont apporté un élément qui nuance l'impact de la caractéristique tridimensionnelle du test en montrant qu'une tâche de rotation en deux dimensions pouvait conduire également à une grande taille d'effet ($d = 1.1$), du moment que la tâche était difficile (stimuli et procédure de comparaison plus complexes). C'est donc la *difficulté* de

6. Ce sont en fait Shepard et Metzler qui ont créé à l'origine ces items (figures abstraites constituées de cubes), repris ensuite par Vandenberg et Kuse dans le MRT.

se représenter et de manipuler mentalement une information tridimensionnelle qui semble déterminant dans l'expression de la différence de sexe, et non la nature — tridimensionnelle ou non — de cette information.

La méta-analyse de Voyer et coll. (1995) confirme l'importance particulière des tests de RM en 3D pour montrer une différence de sexe. Plutôt que de définir *a priori* la notion d'aptitude visuo-spatiale, ces auteurs ont posé le principe que chaque test visuo-spatial définissait opérationnellement une composante précise des aptitudes visuo-spatiales, c'est à dire que chaque test visuo-spatial est selon eux une mesure d'une capacité visuo-spatiale particulière. En effet ils ont pu observer qu'à chaque test était associé une taille d'effet précise, dont la valeur était homogène entre les différentes études concernant des tests identiques. Aussi, Voyer et coll. ont conclu que les tests de RM de Vandenberg et Kuse (3D) montraient la taille d'effet la plus importante ($d = .70$ ou $d = .94$ selon le système de cotation).

1.3.2 Capacité en mémoire de travail

Comment expliquer que la différence de sexe dans le domaine visuo-spatial s'exprime aussi singulièrement en RM ? Mettre à jour les mécanismes cognitifs sous-jacents de la RM peut apporter un éclairage à cette question. Dans son ouvrage de référence, Halpern (2000) attire l'attention sur le fait que généralement les modèles de l'architecture cognitive s'organisent distinctement autour des modalités verbales et spatiales, les différences de capacité observées entre les sexes signeraient donc une différence inscrite au niveau de ces composantes. Par exemple, Baddeley (1986) et Baddeley (2002) distinguent clairement un sous-système de traitement de l'information verbale (boucle phonologique) et un système de traitement visuo-spatial (calepin visuo-spatial). Shah et Miyake (1996) ont montré que ces composantes étaient relativement indépendantes en mesurant les performances obtenues à des tâches conçues pour s'adresser spécifiquement à l'une ou l'autre

composante : les ressources cognitives pouvaient leur être affectées indépendamment selon que l'information à traiter est verbale ou visuo-spatiale. En outre, certains auteurs ont pu montrer que la mobilisation de l'une ou l'autre composante était associée à l'activation de structures cérébrales spécifiques (A. M. Owen et coll., 1999). Ces éléments permettent d'avancer diverses hypothèses explicatives de la différence de sexe en RM. Par exemple, ces différences correspondraient à une différence d'efficacité de ces composantes due à des différences anatomo-fonctionnelles dans les structures cérébrales associées à ces dernières. Ces différences anatomo-fonctionnelles pourraient être un effet de la division sexuelle du travail et des tâches au cours de l'évolution humaine (Silverman, Eals, Barkow, Cosmides, & Tooby, 1992 ; Geary, 1995). Elles pourraient être aussi fondées sur les expériences socialisatrices propres à chaque sexe au cours du développement.

Néanmoins, cette différenciation en fonction de la nature de l'information à traiter est de plus en plus remise en cause. Par exemple, selon Halpern et Wright (1996) — voir également Halpern (2000) —, on constate que certaines tâches mettant en jeu du matériel verbal comme les problèmes d'analogies ou de logique montrent une supériorité des hommes (Meehan, 1984 ; Halpern & Wright, 1996). À l'opposé, des tâches visuo-spatiales seraient mieux réussies par les femmes (dans certaines conditions), par exemple des tâches de mémoire visuelle (Eals & Silverman, 1994 ; Harshman, Hampson, & Berenbaum, 1983 ; James & Kimura, 1997 ; Voyer, Postma, Brake, & Imperato-McGinley, 2007). Or les tâches d'analogie et celles de RM, où les hommes excellent, sont des tâches qui nécessitent un maintien et un traitement actif (manipulation) de l'information — notamment en mémoire de travail. À l'inverse, les femmes réussissent mieux les tests de fluence verbale et de mémorisation de la localisation d'objets dans l'espace, tâches qui nécessitent plutôt le rappel d'information stockée en mémoire à long terme — un traitement de l'information qui par contraste apparaît plus « statique ».

L'importante différence observée en RM serait donc due à une infériorité des femmes à manipuler une information complexe en mémoire de travail. Cornoldi et Vecchi (2003)

ont comparé les performances d'individus des deux sexes à un ensemble de tests destinés à évaluer la mémoire de travail visuo spatiale — certains tests évaluaient les capacités de maintien de l'information, d'autres celles de traitement actif de cette information — et ont conclu que les différences s'observaient essentiellement lors du traitement actif de l'information visuo-spatiale. Si ces auteurs n'ont pas étudié la RM en elle-même, nous pouvons admettre que les problèmes de RM demandent un traitement intensif de l'information et d'ailleurs constituent une mesure de la capacité de la mémoire de travail visuo-spatiale. Un élément allant dans ce sens est que des problèmes de RM sont intégrés à une batterie de tests utilisés par certains chercheurs pour évaluer la composante visuo-spatiale de la mémoire de travail (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). D'autre part, du point de vue anatomo-fonctionnel, il semblerait que la rotation mentale active les mêmes aires cérébrales que les tâches de mémoire de travail visuo-spatiale (M. S. Cohen et coll., 1996 ; Curtis, 2006).

Néanmoins, les études qui ont tenté de mesurer conjointement les différences de sexes en RM et en mémoire de travail n'ont pas été très concluantes. Loring-Meier et Halpern (1999) ont mesuré les performances à différentes tâches visuo-spatiales (visualisation, rappel de formes simples, analyse de scène visuelle, rotation mentale) sensées mobiliser spécifiquement les composantes de maintien ou de manipulation de l'information de la mémoire de travail visuo-spatiale. L'hypothèse de ces auteurs était que les hommes se montreraient supérieurs dans les tâches nécessitant de manipuler l'information, tandis que les femmes seraient supérieures dans les tâches de maintien. Or Loring-Meier et Halpern n'ont observé aucun effet de sexe dans aucune tâche au niveau de la justesse des réponses (le test de RM était en 2D). Par contre, concernant les temps de réponse, un effet de sexe a été rapporté pour tous les types de tâches. L'effet de sexe en question n'était donc vraisemblablement pas spécifique à l'une ou l'autre composante de la mémoire de travail. Dans une autre étude, Robert et Savoie (2006), également inspirées par les propositions d'Halpern (2000), ont mesuré la capacité en mémoire de travail d'hommes et de femmes au niveau de la composante verbale (*digit span*, *reading span*, *verbal span*, *speaking span*)

et de la composante spatiale (matrice de Corsi, rappel de position, *spatial span*), tout en leur soumettant également des problèmes de RM et de fluence verbale. Si ces deux types de problèmes ont révélé les différences observées classiquement, aucune différence de sexe n'est apparue par contre concernant la capacité en mémoire de travail. Cependant ces auteurs n'ont mesuré que la justesse des réponses (comme c'est classiquement le cas lorsque l'on évalue les capacités en mémoire de travail), et non la vitesse de traitement, alors que les précédents résultats de Loring-Meier et Halpern (1999) suggéraient que l'effet de sexe était localisé au niveau des temps de réponse. Par ailleurs, Richardson et Vecchi (2002) ont observé que chez les jeunes adultes, l'effet de la charge cognitive due à la complexité d'une tâche de manipulation de l'information visuo-spatiale (telle que la RM) n'affectait que les temps de réponse. Il semble donc que si la différence de sexe intervient au niveau du traitement actif de l'information, c'est de manière relativement subtile, qui nécessite pour être mise à jour une méthodologie adaptée.

1.3.3 Stratégies

Stratégies de traitement de l'information

Un autre facteur envisagé pour expliquer la différence de sexe est le type de stratégie adoptée par les sujets : en effet, différentes stratégies sont envisageables dans le traitement de l'information visuo-spatiale, y compris pour les problèmes de RM (Kosslyn, 1981 ; Eme & Marquer, 1998 ; Francis & Irwin, 1997 ; Shiina, Saito, & Suzuki, 1997). Les différentes stratégies envisageables peuvent se révéler plus ou moins efficaces, facteur d'autant plus sensible dans le cas des tests difficiles (Cochran & Wheatley, 1988 ; Glück & Fitting, 2003 ; Just & Carpenter, 1985 ; Kanamori & Yagi, 2002 ; McGee, 1978). Précisons ici que cette hypothèse est compatible avec celle portant sur la capacité en mémoire de travail, tant il est vrai dans ce domaine de recherche que les différentes hypothèses formulées sont rarement mutuellement exclusives. Ainsi les femmes adopteraient une stratégie leur per-

mettant de s'adapter aux difficultés rencontrées pour manipuler activement l'information visuo-spatiale, tandis que l'aisance des hommes à ce niveau leur permettrait d'employer une stratégie, en l'occurrence analogique, plus efficace.

L'idée communément partagée aujourd'hui est que les hommes privilégieraient cette stratégie *analogique*, qui correspond à la manière dont la RM est généralement appréhendée, c'est-à-dire par la rotation d'une représentation mentale *analogue* à celle d'un objet réel. Les femmes quant à elles adopteraient plutôt une stratégie de type *analytique*, qui précisément s'appuie sur la *décomposition* de la structure de l'objet (Butler et coll., 2006 ; Clements-Stephens, Rimrodt, & Cutting, 2009 ; Moody, 1998). Par exemple, une fois l'objet décomposé en sous-parties, les opérations de RM seraient simplifiées car elles seraient alors appliquées sur ces sous-parties, par définition moins complexes que l'objet dans sa totalité. Une autre option consisterait à ne se concentrer que sur certains éléments caractéristiques des objets lors de la comparaison entre le modèle et la cible (orientation de certains éléments, intersections entre les éléments). De fait, une stratégie analytique est moins rapide, voire efficace, que la RM car plusieurs étapes (au moins autant que le nombre de sous-parties à traiter) sont alors nécessaires pour sélectionner la réponse. Par contre cette stratégie a l'avantage d'être toujours accessible lorsque les ressources cognitives disponibles sont limitées, car la manipulation en une opération unique de l'objet dans sa totalité consomme évidemment plus de ressource qu'opérer séquentiellement sur chaque sous-partie.

Plusieurs résultats expérimentaux étayaient cette approche, bien que souvent en demi-teinte. Alington, Leaf, et Monaghan (1992) ont montré qu'en facilitant la décomposition des objets représentés dans le MRT en colorant différemment chaque sous-parties de ces objets, on augmentait les performances à ce test, cette augmentation concernant majoritairement les femmes. Ils n'ont par contre pu répliquer ce résultat en apposant des signes distinctifs sur les sous-parties (hachures, pointillés, etc.) qui auraient dûs, pour les mêmes raisons, avoir un effet facilitateur. Peters et coll. (1995), à l'aide d'un questionnaire pro-

posé après un test de RM 3D, ont rapporté que les hommes sont plus nombreux à déclarer faire appel à une stratégie de rotation analogique, alors que, proportionnellement, plus de femmes déclaraient utiliser une stratégie analytique ou s'aider de mouvements des mains, ou du stylo (une autre stratégie pour s'adapter à une charge cognitive trop importante). Néanmoins, la stratégie analogique restait majoritaire chez les femmes.

Privilégiant une approche indirecte, Geiser, Lehmann, et Eid (2006) ont procédé à une analyse fine des erreurs observées au MRT de Vandenberg et Kuse pour en déduire les stratégies employées. Grâce à une analyse par classes latentes d'un très grand nombre de protocoles individuels ($N = 1695$), ils ont pu établir différents profils de réponses au MRT, a priori dûs à la mise en œuvre de différentes stratégies. Les femmes étaient sur-représentées dans la classe qui regroupait les sujets montrant une trop grande lenteur dans la mise en œuvre de la stratégie de RM (analogique). Néanmoins le nombre de femmes et d'hommes dans la classe correspondant au recours à une stratégie analytique n'était pas très différent. Il semble donc bien que les femmes présentent une certaine difficulté à appliquer une stratégie qui correspond à une manipulation active de l'information visuo-spatiale, mais n'appliquent pas pour autant une stratégie différente.

Par ailleurs, beaucoup de chercheurs qui étudient la différence de sexe en RM à l'aide des techniques d'imagerie cérébrale (IRMf essentiellement) n'hésitent pas à interpréter les différences d'activation cérébrale observées au cours de la RM comme témoignant de l'application de stratégies différentes (nous présentons ces études plus bas, Butler et coll., 2006 ; Clements-Stephens et coll., 2009 ; Thomsen et coll., 2000 ; Weiss et coll., 2003). Signalons également qu'un nombre important d'études se sont focalisées sur les choix stratégiques dans le domaine visuo-spatial, mais sur des questions plus spécifiques qui sortent du cadre du présent travail, notamment au niveau des capacités d'orientation (navigation) et de la reconnaissance de scènes visuelles (Astur, Ortiz, & Sutherland, 1998 ; Castelli, Latini Corazzini, & Carlo Geminiani, 2008 ; Dabbs, Chang, Strong, & Milun, 1998 ; Jordan, Schadow, Wüstenberg, Heinze, & Jäncke, 2004 ; Schmitzer-Torbert, 2007). Ces études

semblent quant à elles conclure plus nettement sur l'utilisation de stratégies différentes pour se repérer dans l'espace. Enfin, dans un domaine autre que l'étude des différences de sexe mais avec lequel des analogies sont possibles (vieillissement cognitif), il a été démontré que les personnes âgées adoptent vraisemblablement une stratégie analytique afin de s'adapter à une charge cognitive trop importante (relativement aux capacités du sujet, Dror, Schmitz-Williams, & Smith, 2005). On peut donc envisager que les femmes mettent en œuvre une stratégie non-optimale car l'opération de RM semble conduire à une charge cognitive trop lourde pour elles. Néanmoins, cela n'a pour le moment pas encore été démontré formellement. Le type de stratégie employé semble expliquer en partie l'infériorité des femmes en RM, mais il ne semble pas pour autant être le facteur le plus critique.

Stratégies de sélection de la réponse

Par « stratégies de sélection de la réponse », nous faisons référence à un comportement des sujets qui n'est pas directement lié à l'efficacité avec laquelle ils traitent l'information visuo-spatiale, mais plutôt à une stratégie qui guide la sélection de la réponse en fonction des contraintes du test. Plus précisément, cette idée, qui a bénéficié d'une attention particulière dans la littérature, est que si les femmes sont conscientes de leurs mauvaises aptitudes spatiales (ou y croient simplement), elles peuvent adopter une attitude très conservatrice (prudence) à l'égard de leurs réponses, et ne sélectionner une réponse que lorsqu'elles en sont sûres. Il en résulterait que les erreurs des femmes seraient surtout des erreurs d'omission, d'où un nombre de tentatives typiquement inférieur à celui des hommes, qui pour leur part hésiteraient moins à donner une réponse même lorsqu'ils n'en sont pas sûrs (*guessing*). Par ailleurs, cette plus grande prudence conduirait les femmes à mûrir plus profondément leurs réponses, d'où une latence plus longue, qui leur serait défavorable car les tests montrant une différence de sexe sont typiquement administrés en temps limité.

Goldstein, Haldane, et Mitchell (1990) ont proposé de tester l'hypothèse des facteurs de performances en appliquant une cotation différente au MRT. Plutôt que de comptabiliser les réponses correctes, ces auteurs ont calculé le *ratio* entre le nombre de problèmes correctement résolus et le nombre de problèmes tentés. Si les femmes, plus prudentes, ne proposent que les réponses dont elles sont sûres, alors il est possible qu'*en proportion* elles réussissent le test aussi bien que les hommes, qui, plus confiants, répondraient à un nombre plus important de problèmes, mais plus souvent faux également. Les résultats de Goldstein et coll. confirment bien cette hypothèse : la supériorité des hommes au MRT disparaît si les scores sont calculés en fonction des ratios *réussis / tentés*. Ces auteurs observent également une annulation de l'effet de sexe lorsque le test était administré sans pression temporelle⁷. Il semble donc que les capacités en RM ne sont pas le seul facteur à considérer pour expliquer la différence de sexe dans de ce domaine.

Néanmoins d'autres résultats ont amené à relativiser l'importance de ce type de stratégie. Stumpf (1993) a essayé d'étendre le résultat de Goldstein et coll. (1990) à 15 tests visuo-spatiaux et a observé que prendre en compte les ratios *réussis / tentés* ne conduisait pas à une réduction de l'effet de sexe pour tous les tests. Par exemple, il rapporte pour le MRT une réduction de l'effet de $d = .74$ à $d = .48$ (et non une annulation), mais aucune réduction ($d = .72$ vs. $d = .68$) pour le test de RM d'Eliot-Price. Or ce dernier test est similaire au MRT, sauf qu'il faut pour chaque problème sélectionner une proposition parmi quatre plutôt que deux parmi quatre (MRT). Prinz et Freeman (1995) ont proposé une tâche de RM où les deux objets à comparer apparaissaient l'un après l'autre sur un écran d'ordinateur, avec une mesure précise des temps de réponse, afin de déterminer si les femmes avaient tendances à privilégier la justesses des réponse par rapport à la rapidité, à l'inverse des hommes. Les résultats ont montré que si l'effet de sexe était réduit lorsque le ratio *réussis / tentés* était calculé (de $d = .84$ à $d = .45$), cet effet n'était pas annulé pour autant. Par ailleurs, les femmes étaient bien plus lentes que les hommes sans pour autant faire moins d'erreurs, excepté pour les problèmes les plus difficiles. L'effet des stratégies

7. Goldstein et coll. (1990) avaient par ailleurs introduit le score obtenu au test standardisé en mathématique SAT-M en covarié.

semble donc assez volatile, très dépendants du « contexte », c'est-à-dire du type de test, du type de problème, etc.

D'autres auteurs encore n'ont décelé aucun effet ni du système de cotation, ni de la pression temporelle. Delgado et Prieto (1996) en manipulant les instructions — pour mettre l'accent soit sur la nécessité de répondre vite soit sur celle de répondre correctement — et la pression temporelle n'ont pas pu répliquer les résultats de Goldstein et coll. (1990). Dans toutes les conditions les hommes avaient des résultats supérieurs et aucun des deux groupes de sexe ne s'est distingué par un excès (ou un défaut) de prudence dans ses réponses. Masters (1998) obtient des résultats sensiblement identiques. Peters (2005) n'observe aucun effet de la pression temporelle alors même que l'effet du sexe est plus important pour les problèmes situés à la fin de test : les effets sont donc dépendants du temps passé sur la tâche, mais le lien avec la stratégie de sélection de la réponse n'est pas évident.

Enfin, Voyer (1997) a abordé l'hypothèse des stratégies en établissant une typologie des erreurs au MRT. Il a cherché à isoler des profils de réponse typiques d'erreurs d'omissions, révélatrices d'une stratégie plus conservatrice, a priori majoritaire chez les femmes. Les résultats n'ont pas confirmé cette hypothèse, de même que la manipulation de la pression temporelle n'a pas affecté les écarts entre les groupes de sexe. Par contre le calcul des scores au MRT à partir du ratio *réussis/tentés* réduisait la différence de sexe. Par la suite, Voyer et Saunders (2004) ont montré à l'aide d'une analyse factorielle des profils de réponse qu'il existait bien une catégorie distincte formée par les individus réalisant des erreurs d'omission, mais aucun des deux sexes n'était sur-représenté dans cette catégorie. Dans une dernière étude où ils ont systématiquement manipulé la pression temporelle, Voyer, Rodgers, et McCormick (2004) n'ont pas observé un effet de la pression temporelle sur l'expression des différences de sexe, mais ils ont pu observer que plus les femmes disposaient de temps, plus elles montraient des profils de réponse correspondant à des erreurs d'omission.

Les études portant sur les stratégies de sélection de la réponse livrent donc des résultats en demi-teinte sur la validité de cette hypothèse. Le phénomène semble assez volatil, et dépendre fortement des conditions de passation des tests (instructions, pression temporelle, type de test, difficulté, etc.), ce qui est certes conforme à l'hypothèse. Mais la différence de sexe, elle, est observée beaucoup plus systématiquement et semble donc témoigner davantage d'une réelle infériorité des femmes dans le domaine visuo-spatial, que de l'action de facteurs extérieurs. Les stratégies de sélection semblent donc jouer seulement un rôle partiel dans l'expression des différences de sexe. En outre, cette hypothèse est difficile à tester : la lenteur des femmes est-elle due à de plus grandes hésitations du fait d'un manque de confiance ou à une plus grande difficulté à manipuler l'information visuo-spatiale ? Sans compter que les deux phénomènes, non exclusifs, peuvent très bien s'additionner.

1.4 Facteurs biologiques et sociaux

Nous avons vu précédemment quels étaient les facteurs cognitifs qui ont été majoritairement envisagés par la recherche sur les différences de sexe pour expliquer ces dernières en RM. Ce domaine de recherche est traversé par une question plus fondamentale : ces différences sont-elles essentielles — c'est-à-dire un fait de nature, inné, biologiquement inscrit — ou construites — c'est-à-dire culturelles, acquises au cours du développement de l'individu dans un environnement social donné ? Bien sûr, aujourd'hui, cette question n'est plus abordée dans le sens d'une opposition « frontale » entre nature et culture. Un consensus émerge depuis quelques années autour de modèles qu'Halpern (2000) qualifie de « psycho-bio-sociaux », c'est-à-dire qui décrivent l'interaction entre les déterminants biologiques et sociaux du comportement. Par exemple Casey, Brabeck, et Nuttall (1995) évoquent l'interaction entre les facteurs sociaux et les facteurs biologiques à l'aide de la métaphore de la brindille qui, tordue dans une direction donnée, se développe en une branche qui pointe fermement dans cette direction (modèle « bent-twist », voir aussi Casey,

1996b, 1996a). La différence de sexe due à des facteurs biologiques, faible lorsque les sujets sont très jeunes, serait amplifiée au cours de la socialisation des individus, qui, en fonction de leur sexe, ne seraient pas encouragés et entraînés à développer les mêmes compétences.

Nous allons donc dans cette dernière section présenter les résultats qui montrent l'implications des facteurs biologiques, puis sociaux, dans la différence de sexe en RM. Concernant les facteurs sociaux, nous concluons sur des résultats qui suggèrent la nécessité de rompre avec l'approche classique accordant à l'internalisation de certains stéréotypes négatifs un rôle clef.

1.4.1 Le substrat biologique des opérations de RM

Une des théories de l'origine biologique des différences de sexe les plus populaires aujourd'hui adopte une perspective évolutionniste (Eals & Silverman, 1994 ; Geary, 1995, 1998 ; Kimura, 2004). La division du travail au cours de l'évolution humaine, notamment dans les sociétés préhistoriques de chasseurs-cueilleurs, associée au mécanisme de sélection sexuelle, aurait entraîné une spécialisation des structures cérébrales masculines dans le traitement de l'information visuo-spatiale. Les hommes les plus à même de s'assurer une descendance auraient été les meilleurs chasseurs, c'est-à-dire les hommes les plus capables de se repérer dans l'espace, ce qui nécessite de bonnes capacités en RM. Une telle pression sélective aurait abouti chez les hommes à une supériorité génétiquement inscrite dans le domaine visuo-spatial. Cette théorie explique également la meilleure mémoire visuelle des femmes, et donc leur propension à utiliser une stratégie analytique, par des activités (soin du foyer, cueillette) les conduisant à être attentives aux détails et à mémoriser des indices spatiaux spécifiques pour se repérer dans un environnement plus limité et plus stable que celui des hommes.

Cette explication évolutionniste des différences de sexe repose toutefois sur des hypothèses aujourd'hui encore discutées. Par exemple la division du travail était semble-t-il

beaucoup plus complexe qu'une simple dichotomie entre chasse et cueillette qui par sa grossièreté renvoie plutôt à une image d'Épinal, cette dernière étant mise à mal par de nombreuses observations anthropologiques (Testart, 1986 ; L. Owen, 2005). En outre, les chercheurs réévaluent sans cesse l'estimation, toujours plus tardive, du moment auquel cette division est sensée survenir dans l'histoire humaine : dernièrement, S. L. Kuhn et Stiner (2006) ont présenté une théorie selon laquelle la division sexuelle du travail ne se serait généralisée qu'au début du paléolithique supérieur, il y a 30 000 ans environ⁸ (on estime qu'*Homo sapiens* est apparu il y a 200 000 ans). Or les auteurs « évolutionnistes » tels Tooby et DeVore (1987) envisagent une division du travail beaucoup plus précoce : c'est selon eux par cette division que les premiers *Hominines* se seraient distingués des grands singes (5 millions d'années), ou tout au moins aurait-elle concerné *Homo erectus* (1 million d'années) dont l'alimentation était pour une bonne part carnée, preuve que cette espèce consacrait une bonne part de son activité à la chasse.

Malgré ces réserves, on ne peut ignorer le substrat biologique de l'activité cognitive, et les résultats qui dans ce domaine montrent une différence d'activité cérébrale associées aux différences de sexe observées dans le domaine visuo-spatial en général, et de la RM en particulier. Par un souci d'économie nous limitons notre exposé aux résultats concernant les facteurs neuro-anatomiques, sans aborder le rôle des hormones sexuelles, malgré la richesse de ce champ de recherche. Signalons simplement à ce sujet qu'il est établi pour le modèle animal que le niveau de testostérone a un effet positif sur les performances visuo-spatiales, tandis que le niveau d'œstrogène a lui un effet négatif, les performances des femelles étant maximum au moment de leur cycle menstruel où ce niveau est minimum (Healy, Braham, & Braithwaite, 1996). Malgré des données qui ont été longtemps difficiles à interpréter chez le sujet humain (McKeever, 1995), des résultats allant dans ce sens ont depuis été publiés (Hausmann et coll., 2000). Nous avons choisi de ne pas détailler les résultats concernant les hormones car la manière dont leur action affecte des mécanismes

8. Selon S. L. Kuhn et Stiner, la division du travail aurait garanti une bonne efficacité économique, laquelle expliquerait l'explosion démographique et la colonisation de l'Eurasie par *Homo sapiens* à cette date, parallèlement au déclin d'*Homo neanderthalensis* qui n'avait pas adopté cette organisation.

cognitifs est moins évidente que les relations aujourd’hui connues entre ces mécanismes et les structures neuro-anatomiques. Notons qu’il semble que les œstrogènes ont un effet négatif au niveau de la mémoire de travail (Duff & Hampson, 2000), et surtout au niveau de l’activité des structures cérébrales préfrontales (Keenan, Ezzat, Ginsburg, & Moore, 2001 ; Galea et coll., 2008), impliquées comme nous allons le voir dans la RM.

Latéralisation fonctionnelle.

L’une des hypothèses les plus abondamment documentées concernant la différence d’organisation cérébrale supposée entre les sexes porte sur la latéralisation fonctionnelle. L’organisation anatomo-fonctionnelle des femmes serait plus symétrique (bilatéralisation) que celle des hommes, plutôt asymétrique (latéralisation). Chez ces derniers, des structures de l’hémisphère gauche seraient plus fortement mobilisées pour traiter l’information verbale, tandis que le traitement de l’information visuo-spatiale concernerait majoritairement des structures de l’hémisphère droit. Les premières observations en ce sens étaient des constats cliniques : des lésions cérébrales comparables — et localisées dans un hémisphère donné — ne provoquent pas toujours les mêmes déficits fonctionnels chez les hommes et les femmes (Springer & Deutsch, 1998). Chez les hommes, une lésion située sur l’hémisphère gauche est susceptible d’entraîner un déficit des capacités verbales tandis qu’une lésion de l’hémisphère droit entraîne un déficit dans les capacités visuo-spatiales. De telles observations sont beaucoup plus rares chez les femmes, les aphasies consécutives à une lésion de l’hémisphère gauche survenant par exemple trois fois moins souvent chez ces dernières (J. McGlone, 1978). De plus, la latéralisation fonctionnelle semble avoir été un phénomène essentiel au cours de l’évolution du genre *Homo*, témoignant de manière singulière de la sophistication croissante des capacités cognitives humaines (Corballis, 1989). Cette perspective est compatible avec l’hypothèse évolutionniste sur les différences de sexe évoquées précédemment.

Expérimentalement, chez le sujet sain, la latéralisation fonctionnelle pour le traitement de l'information visuo-spatiale peut être étudiée par le recours aux techniques d'imagerie cérébrale (EEG, IRMf, PET-scan, etc.), mais aussi par une méthode comportementale bien moins lourde : la présentation des stimuli visuels en champs divisés. En effet, l'information visuelle en provenance de la zone latérale périphérique du champ de chaque œil est dans un premier temps traité exclusivement par les structures de l'hémisphère controlatéral. Si les stimuli présentés dans l'hémichamp gauche sont traités plus efficacement que ceux présentés dans l'hémichamp droit, on pourra conclure à un traitement plus efficace de l'information par les structures de l'hémisphère droit. Une autre méthode consiste à comparer des individus qui diffèrent du point de vue de la préférence manuelle (droitiers vs. gauchers).

Croisant les résultats de plusieurs études qui mettaient en œuvres l'ensemble des méthodes présentées précédemment, J. McGlone (1980) conclut à la validité de l'hypothèse de la latéralisation. Mais les conclusions de Fairweather et Beaumont (1982), se focalisant sur les expériences en champ divisé, vont moins nettement dans ce sens. Effectivement, la synthèse des données dans ce domaine n'est ni aisée, ni concluante, mais « chaotiques » selon McKeever (1995). Dans une méticuleuse méta-analyse incluant le calcul des tailles d'effet, Voyer (1996) est parvenu à un résultat en apparence paradoxal, mais qui peut précisément expliquer ces résultats contradictoires. Alors que les différences de sexe les plus importantes sont observées dans le domaine visuo-spatial, la spécialisation hémisphérique s'y révélerait moins importante que dans le cas des tâches verbales où la différence de sexe est pourtant modérée. Ce résultat surprenant est confirmé par Corballis (1997). Comment donc la latéralisation pourrait-elle expliquer la différence de sexe si les effets de la latéralisation sont modérés lorsque la différence de sexe est importante, et vice-versa ?

En fait, d'autres résultats ont suggéré que de nombreux facteurs intervenaient dans les situations qui permettaient de mettre à jour une latéralisation fonctionnelle. Alors que, sur une tâche de RM présentée en champs divisés, Rilea, Roskos-Ewoldsen, et Boles (2004)

n'ont décelé aucun effet de sexe, ni d'effet de latéralisation. Pourtant, Siegel-Hinson et McKeever (2002) avaient réussi auparavant à montrer que la latéralisation des hommes permettait d'expliquer leur supériorité dans le domaine visuo-spatial. L'étude de Siegel-Hinson et McKeever se distinguait sur deux points. Le paradigme de champ divisé nécessite de présenter les stimuli très rapidement (de l'ordre de 200 ms) afin de s'assurer que ceux-ci sont effectivement projetés dans la périphérie de l'hémichamp du sujet, sans que ce dernier n'ait le temps de déclencher une saccade oculaire qui centrerait le stimuli dans le champ. Vu ce laps de temps très court, il faut donc que les problèmes proposés restent simples. Rilea et coll. avaient par exemple choisi une RM en deux dimensions. Or les problèmes de RM en 2D ne permettent pas de mettre en évidence un effet de sexe, ce qui expliquerait pourquoi ces auteurs n'ont pu déceler aucune différence de sexe. Siegel-Hinson et McKeever (2002) ont choisi pour leur part un test où ils demandaient de comparer des cubes dont les faces portaient des signes distinctifs. Ces cubes étaient vus depuis deux points de vue différents, ce qui faisait intervenir une forte composante de RM dans ce test, en 3D, mais relativement simple. Par ailleurs ce test ne comportait qu'un nombre relativement restreint de problèmes (60), car l'hypothèse de travail de ces auteurs était également que si la tâche durait trop longtemps, les sujets pouvaient mettre en œuvre des stratégies d'ordre « symbolique » ou « analytique » qui auraient mobilisé l'hémisphère gauche.

Il apparait donc que l'effet de latéralisation est sensible au type de tâche à résoudre et de la stratégie mise en œuvre par le sujet (ces deux facteurs étant souvent liés, on s'en doute). Si Vogel, Bowers, et Vogel (2003) dans une méta-analyse ont pu conclure que les hommes étaient plus latéralisés que les femmes, ces auteurs ont également constaté que la latéralisation dépendait aussi du type de tâche. Par exemple la résolution des tests de visualisation spatiale n'était associés à aucun effet de latéralisation. Parallèlement, utilisant des techniques d'électro-encéphalographie (EEG), Johnson, McKenzie, et Hamm (2002) ont observé une latéralisation spécifique chez les hommes lors d'une tâche de RM en 2D (lettres), résultat confirmé avec la même technique par Roberts et Bell (2003)

qui, par contre, ont constaté pour la RM 3D une activation du lobe pariétal droit pour les deux sexes : la RM ne permettrait donc pas d'observer d'effet de la latéralisation. Enfin, Tomasino et Rumiati (2004) ont montré que des patients souffrant de lésions à l'hémisphère gauche rencontraient des difficultés face à un test de RM 3D, tout comme ceux souffrant de lésions à droite, et qu'en fait le type de lésion affectait le type de stratégie mise en œuvre.

Finalement, la différence de latéralisation observée entre les sexes reflète-t-elle une différence d'organisation cérébrale, conduisant chaque sexe à réaliser les RM de manière spécifique (stratégies), ou l'utilisation de stratégies différentes pour chaque sexe, conduisant à des activations différentes ? Un des problèmes majeur dans ce domaine est la difficulté à sortir d'une causalité réversible.

Imagerie cérébrale

Thomsen et coll. (2000) ont mené la première étude IRMf conçue explicitement pour tester avec plus de précision l'hypothèse de la latéralisation. Auparavant, en IRMf, seuls Tagaris et coll. (1997) avaient comparé les activations cérébrales des deux sexes au cours d'un test de RM, mais ils avaient conclu à une activation du cortex pariétal supérieur dans les deux groupes. Thomsen et coll. quant à eux ont également constaté dans les deux groupes une activité bilatérale dans le cortex pariétal, mais cette activité était plus importante dans l'hémisphère droit chez les hommes, et chez les femmes, une activité du gyrus frontal inférieur droit (zone du cortex préfrontal). La différence d'activation entre les groupes de sexe ne semble pas tant concerner la latéralisation qu'un axe antéro-postérieur au sein de l'hémisphère droit. Thomsen et coll. interprètent ces activations comme le signe d'une différence dans les stratégies employées. Pour notre part, nous signalons qu'il a été montré récemment que le gyrus frontal inférieur, dont l'activation est observée ici chez les femmes, est mobilisé dans les situations de prises de décision chez les sujets qui montrent une aversion au risque (Christopoulos et coll., 2009). Nous pouvons supposer que d'autres

processus que ceux liés au traitement de l'information spatiale peuvent donc intervenir au cours du test.

Autre point troublant, Thomsen et coll. (2000) n'ont observé aucun effet de sexe, sinon tendanciel. D'ailleurs, Unterrainer, Wranek, Staffen, Gruber, et Ladurner (2000) partant du constat que de nombreuses études en imagerie ne parvenaient pas à mettre en évidence de différences de sexe concernant la latéralisation, ont émis l'hypothèse que la latéralisation du traitement au cours d'une tâche pouvait non pas être liée au sexe, mais plutôt à la qualité de la performance. Mettant en œuvre une variante de la technologie TEP⁹, ces auteurs n'ont pu mettre en évidence de différences significatives entre l'activité cérébrale des hommes et des femmes à un test de RM. Par contre, lorsque les sujets étaient comparés en fonction de leur niveau de performance, ceux qui réussissaient bien montraient une asymétrie statistiquement décelable — activation de la zone parieto-temporo-occipitale droite seulement —, tandis que ceux qui réussissaient mal présentaient une activation bilatérale de cette même zone. Il semble donc selon ces résultats que les différences d'activations reflètent plus des différences dans la réalisation du traitement visuo-spatial, plus ou moins efficace, qu'un « cablage neuronal » spécifique à l'un ou l'autre sexe.

Néanmoins, Jordan et coll. (2002) ; Weiss et coll. (2003), prenant la précaution de ne tester que des hommes et des femmes avec de bonnes capacités en RM, ont mesuré des différences d'activation compatibles entre elles et avec celles déjà rapportées par Thomsen et coll. (2000) : essentiellement une activation bilatérale des lobes pariétaux — inférieurs pour les hommes, et supérieurs pour les femmes. De plus, les femmes présentaient une activité supérieure des lobes frontaux droits. Les activations étaient pour l'essentiel bilatérales (sauf au niveau frontal), ce qui contredit les résultats d'Unterrainer et coll.. Ces divergences sont peut-être dues au fait que ces derniers auteurs n'avaient pas utilisé un test de RM strictement, mais une tâche avec simplement une forte composante en RM.

9. Tomographie par Emission de Positons.

Quoi qu'il en soit, la différence d'activité cérébrale entre les deux sexes semble plus fine que suggérée par l'hypothèse de la latéralisation. Des facteurs tels que la performance ou la stratégie employée semblent également intervenir, et le lien causal entre ces facteurs et le sexe des sujets reste dans tous les cas à préciser. En outre, l'activation spécifique de la région préfrontale droite des femmes suggère que des processus qui ne sont pas directement liés au traitement de l'information visuo-spatiale peuvent intervenir, cette région étant vraisemblablement liée aux processus de prise de décision, notamment en rapport à soi, aux émotions, à l'évitement du risque et à l'activité de la mémoire de travail (Christopoulos et coll., 2009 ; Kringelbach, 2005 ; LoPresti et coll., 2008). Une récente étude de Wraga et coll. (2007) apporte des éléments dans ce sens. S'inspirant d'une théorie qui fait intervenir les stéréotypes sociaux (nous consacrons notre chapitre 2 à cette théorie), ces auteurs ont manipulé les attentes des sujets (féminins) quant à un test de RM, en le présentant soit comme typiquement réussi ou échoué par les femmes. Cette manipulation, qui fait intervenir un facteur social — des attentes conformes ou non à une croyance socialement partagée d'infériorité des femmes dans le domaine visuo-spatial —, peut provoquer une modification des activations cérébrales observées au cours du test de RM. Les régions cérébrales associées aux émotions (région rostro-ventrale gauche du cortex cingulaire antérieur, gyrus orbital droit) ont alors été activées différemment entre les deux conditions de manipulation de cette croyance sociale.

Ainsi, des facteurs sociaux peuvent intervenir au cours du test de RM de manière jusqu'ici insoupçonnée. Non seulement la manipulation des attentes en fonction d'un stéréotype de genre a un effet sur l'activation des aires cérébrales concernées — ce qui suggère une perméabilité des processus cognitifs à ces facteurs —, mais la performance des femmes s'en est retrouvée également affectée. Dans la condition où les femmes étaient amenées à nourrir des attentes négatives, leur performance a été inférieure à celle des femmes placées dans la condition où des attentes positives leur étaient suggérées. Nous présenterons plus en détail dans le chapitre suivant l'étude de Wraga et coll. (2007), ainsi que la théorie de psychologie sociale qui l'a inspirée, et à laquelle la psychologie cognitive a été jusqu'à pré-

sent relativement hermétique. Mais auparavant nous terminons ce chapitre en présentant brièvement la manière classique dont les facteurs sociaux sont abordés dans la recherche sur la différence de sexe, notamment en RM.

1.4.2 Les facteurs sociaux

Le rôle de la socialisation

Afin d'expliquer les différences de performance en RM par des facteurs sociaux, on avance généralement l'idée qu'il existe une pression sociale — sous la forme d'encouragements, de conventions voire d'interdits — qui orienterait les individus vers des activités différentes (jeux, sport, apprentissages, etc., cf. Blakemore et coll., 2009 ; Cherney & London, 2006), en fonction de leur sexe. Ces activités ne conduiraient pas chaque groupe de sexe à développer les mêmes compétences, dont celles en RM (Halpern, 2000).

Cette approche n'est pas moins compatible avec une hypothèse évolutionniste que celle ayant trait aux facteurs biologiques. En effet, notre espèce forme avec les grands singes une famille où les interactions sociales ont une importance capitale, et il en était évidemment de même tout au long de l'évolution du genre *Homo* depuis des millions d'années. Il est donc fort probable que nous ayons développé une sensibilité particulière au regard d'autrui, et au sens que portent les interactions sociales. Une preuve de cette sensibilité est que les très jeunes enfants élaborent assez tôt des représentations sur les rôles sociaux. Par exemple, D. Kuhn, Nash, et Bruckner (1978) ont montré que dès l'âge de deux ans, un enfant est capable d'attribuer à un personnage des aptitudes et des rôles conformes aux stéréotypes de genre, en fonction du sexe supposé de ce personnage. Dès cet âge les enfants savent donc ce que l'on attend d'une petite fille et d'un petit garçon, et donc d'eux-mêmes.

Par ailleurs, la force de cet apprentissage social vient aussi du fait qu'il est alimenté par de nombreuses sources (parents et autres adultes, autres enfants, médias, etc. cf. par exemple Bem, 1981 ; Benbow, 1988 ; Witt & Burriss, 2000 ; Witt, 2000). Pour donner un exemple précis, Jacobs (1991) a montré que l'adhésion des parents aux stéréotypes de genre influençait leur perception des capacités de leurs enfants en mathématiques (en fonction du sexe de ces derniers), ce qui affectait ensuite la perception des enfants de leurs propres capacités de manière congruente avec le stéréotype. Ainsi les femmes, auxquelles le stéréotype attribue une infériorité en mathématiques, peuvent être conduites à adopter une attitude de désengagement vis-à-vis des mathématiques, ne pas chercher à développer leurs compétences dans ce domaine ou encore décider de ne pas s'engager dans un cursus scolaire ou une carrière professionnelle en rapport avec les mathématiques.

Cette approche classique suppose donc une intervention diffuse des facteurs sociaux, qui, sur le long terme, vont affecter l'identité du sujet que l'on sait particulièrement sensible à la catégorisation de genre (Bem, 1981), et prompt à développer des capacités congruente avec le stéréotype. Les facteurs sociaux sont donc envisagés d'un point de vue « macro-social », ce qui ne va pas sans poser de difficulté pour mettre leur implication en évidence. Cette implication ne peut être démontrée qu'indirectement, en détectant par exemple les éventuelles relations entre les capacités en RM des sujets et leurs croyances (vis à vis de leur compétences en RM), ou leur expérience acquise à travers diverses activités dans le domaine visuo-spatial. Ce sont donc majoritairement des facteurs invoqués, dont les effets sont généralement établis grâce à des analyses corrélationnelles, qui ne permettent pas de conclure quant à la nature causale des relations dans lesquelles ces facteurs sont impliqués.

Internalisation et identité de genre

Une méthode pour étudier l'effet de l'environnement social sur les compétences en RM est de comparer les performances de sujets appartenant à différents environnements

culturels. Selon l'hypothèse de l'internalisation, les sujets ne développeraient pas la même identité de genre et les compétences associées, en ce qui nous concerne, la RM, selon leur environnement social et culturel. Mais les données de la littérature ne confirment pas cette hypothèse : plusieurs études trans-culturelles ont montré que la supériorité des hommes en RM est observée dans différents continents (Jahoda, 1980 ; Peters, Lehmann, Takahira, Takeuchi, & Jordan, 2006). Plus particulièrement, Halpern et Tan (2001) ont testé les capacités en RM d'étudiant(e)s d'une petite ville dans une région rurale de la Turquie, tout en prenant soin de mesurer leur adhésion aux stéréotypes de sexe. Halpern et Tan s'attendaient à observer des comportements associés à des stéréotypes particulièrement conservateurs à l'égard des femmes. Or si effectivement les étudiants des deux sexes croyaient en la supériorité des hommes en sciences, en athlétisme, etc., ils n'envisageaient pas l'existence d'une différence de sexe dans le domaine visuo-spatial. Et malgré cela, les auteurs ont observé la classique différence de sexe au MRT ($d = .58$).

Néanmoins, dans le domaine de la « construction de soi » (*self-construal*, personnalité, valeurs, émotions), Guimond et coll. (2007) rapportent que, de manière surprenante, les différences de sexe les plus importantes sont observées dans plusieurs pays occidentaux, malgré les progrès *a priori* accomplis au niveau de l'égalité des sexes (comparé avec la Malaisie). En fait il apparaît que cette « définition de soi » est sensible aux comparaisons sociales, notamment intergenres, or ces comparaisons interviennent plus souvent dans des sociétés égalitaires que dans des sociétés dont la structure est plus rigide. Ces résultats suggèrent donc, même s'ils ne concernent pas directement les différences en RM, que les mécanismes de régulations faisant intervenir la culture, les représentations sociales, les stéréotypes, etc. sont beaucoup plus complexes et malléables qu'imaginés par l'hypothèse classique d'internalisation.

Une autre méthode, lorsque l'on n'a accès qu'à des individus appartenant à un même environnement culturel, consiste à mesurer l'identité de genre de ces individus. Ainsi Signorella et Jamison (1986) ont pu établir grâce à une méta-analyse que les test visuo-

spatiaux et de mathématiques étaient mieux réussis par les individus dont l'identité de genre comportait de nombreux traits masculins, et ce, particulièrement chez les femmes. Signorella, Jamison, et Krupa (1989) ont confirmé plus précisément ce résultat sur un test de rotation mentale, et selon Hamilton (1995), la réussite au MRT est associée à une identité androgyne (possession de traits masculins et féminins). Les individus avec de bonnes compétences en RM ont donc bien intériorisé un certain nombre de traits de l'identité masculine. Mais cela correspond-il pour autant à une plus grande pratique, ce qui signifie un meilleur entraînement, dans le domaine visuo-spatial ?

Expérience Spatiale

Les garçons, à qui l'on permet plus volontiers des jeux de plein air et des activités sportives, à qui l'on offre des jeux de construction, et que l'on encourage à s'engager dans des cursus scientifiques et techniques, auraient-ils donc plus d'occasions de cultiver leurs aptitudes en RM ? Tout d'abord, Peters et coll. (1995) ont pu vérifier dans un premier temps que la performance au MRT est très sensible à l'entraînement (scores doublés en quatre semaines d'entraînement), et que cet entraînement est surtout bénéfique aux femmes qui suivent un cursus scientifique. Ces dernières semblent d'ailleurs avoir en général de meilleures performances que le reste des femmes, même si au sein des étudiants en sciences une différence de sexe persiste, bien que réduite (Quaiser-Pohl & Lehmann, 2002).

Newcombe, Bandura, et Taylor (1983) ont étudié plus systématiquement la pratique des activités spatiales en fonction du sexe, et ils ont observé que les sujets jugeaient les activités typiquement masculines (sport, bricolage, etc.) plus révélatrices des capacités visuo-spatiales que des activités mixtes ou féminines. Ces auteurs ont également observé une corrélation entre la pratique d'activités jugées « spatiales » et la performance à des tests visuo-spatiaux, particulièrement pour les femmes (cf. également Quaiser-Pohl & Lehmann, 2002). Néanmoins, selon Siegel-Hinson et McKeever (2002), l'expérience spatiale

(mesurée avec le questionnaire de Newcombe et coll.) ne médiatiserait au mieux qu'en partie l'effet de sexe au MRT, alors qu'un facteur biologique comme la latéralisation fonctionnelle en médiatiserait la totalité.

Enfin, un certain nombre d'études se sont récemment focalisées sur les jeux vidéos, une pratique majoritairement masculine. En effet certains jeux peuvent mobiliser assez intensément les capacités visuo-spatiales du joueur, et ce d'autant plus depuis la généralisation des jeux simulant un environnement en 3D. Ainsi plusieurs auteurs ont rapporté un effet positif de la pratique des jeux vidéos lorsqu'elle était imposée et contrôlée expérimentalement, — il s'agissait essentiellement de jeux qui par leur conception entraînaient les sujets à la RM (Cherney, 2008 ; De Lisi & Cammarano, 1996 ; De Lisi & Wolford, 2002). D'autres auteurs se sont intéressés à la pratique spontanée des jeux vidéos. Peters et coll. (1995) observent que si les hommes déclarent jouer plus fréquemment, aucune corrélation entre cette pratique estimée et la performance au MRT n'est constatée. Quaiser-Pohl, Geiser, et Lehmann (2006) ont déterminé à partir d'un échantillon de grande taille ($N = 861$) différents profils de joueurs, il est apparu que la majorité des femmes soit ne jouaient pas, soit préféraient les jeux de réflexion, alors qu'un troisième groupe, outre les deux précédents, pouvait être formé chez les hommes : les utilisateurs de jeux d'actions et de simulation. Si la pratique d'un jeu vidéo n'avait aucun effet sur la performance des femmes au MRT, elle était bénéfique chez les hommes, quel que soit le type de jeu. Ainsi la différence de sexe la plus faible était observé chez les non-joueurs. Il semble donc que l'entraînement permette aux hommes d'exacerber une habileté « naturelle » : dans cette approche le processus de socialisation ne ferait donc qu'intensifier une différence de sexe pré-existante.

1.5 Conclusion

La différence de sexe, lorsqu'elle est testée à l'aide du MRT (rotation en trois dimensions), figure parmi les différences les plus robustes et les plus importantes en taille, à

l'avantage des hommes. Mais constater une différence ne nous renseigne pas sur son origine. De nombreux facteurs, dont l'interconnexion semble par ailleurs assez forte, peuvent jouer un rôle dans l'expression de cette différence. L'infériorité constatée en RM serait en fait un cas particulier d'une difficulté plus générale pour les femmes à manipuler activement l'information visuo-spatiale en mémoire de travail. Cette difficulté semble associée à la mise en œuvre d'une stratégie moins efficace que la stratégie de RM « analogique », qui exige une grande disponibilité des ressources cognitives. Cette hypothèse est compatible avec les observations en imagerie cérébrale (latéralisation fonctionnelle, activations de structures différentes en fonction du sexe). Néanmoins la question du sens de la causalité n'est toujours pas résolue. Est-ce parce que leurs structures cérébrales diffèrent qu'hommes et femmes mettent en œuvre des stratégies différentes ? Ou est-ce qu'un autre facteur intervient qui pousse les femmes à recourir à une stratégie moins efficace, d'où des activations cérébrales différentes ? Nous avons décidé dans cette thèse de nous saisir de cette dernière question et d'interroger la nature sociale d'un tel facteur. Or les résultats classiques concernant l'intervention de facteurs sociaux au cours de la RM sont assez faibles, et peu cohérents. Nous proposons d'aborder cette question de manière radicalement différente, en ne considérant plus les facteurs sociaux dans leur dimension exclusivement « macro-sociale », mais comme le résultat d'une interaction entre les caractéristiques de la situation de test et certaines connaissances stockées en mémoire à long terme en rapport avec le groupe social d'appartenance.

Chapitre 2

Interférence du stéréotype de genre et rotation mentale

2.1 Introduction

2.1.1 Une alternative à l'approche macro-sociale

Nous allons, dans ce chapitre, présenter en détail le paradigme de la menace du stéréotype (*stereotype threat*, Steele, 1997), et certaines données récentes en démontrant la pertinence pour l'étude de la différence de sexe en RM. Dans ce paradigme les stéréotypes occupent une place centrale pour expliquer les différences de performance observées entre les sexes dans plusieurs domaines, notamment les mathématiques, et plus récemment dans celui des rotations mentales. Cette théorie constitue une alternative originale et sérieuse d'une part à l'hypothèse d'une différence de performance fondée biologiquement et d'autre part à l'hypothèse d'internalisation des stéréotypes comme condition nécessaire à leur influence sur le comportement. Enfin elle propose une méthodologie expérimentale rigoureuse pour mettre en évidence l'interférence du stéréotype qu'elle suppose. Avant

de détailler ces différents aspects, il convient de préciser la notion de stéréotype, ce qui permettra de mieux saisir l'originalité de l'hypothèse de Steele.

2.1.2 Les stéréotypes

Selon Leyens, Yzerbyt, et Schadron (1994, p. 24), les stéréotypes constituent des *« croyances partagées concernant les caractéristiques personnelles, généralement des traits de personnalité, mais souvent aussi des comportements, d'un groupe de personnes »*. Toujours selon ces auteurs, *« utiliser un stéréotype donné revient à considérer que tous les membres d'une catégorie [...] partagent les attributs contenus dans le stéréotype »*. Par exemple, « Les femmes sont mauvaises en mathématiques », ou plus directement lié à notre sujet : « Les femmes ne savent pas lire les cartes routières. ». Cette généralisation abusive, qui conduit à prêter un trait particulier à la totalité des membres d'un groupe social est une caractéristique essentielle du stéréotype. Toute femme, aussi forte soit-elle en mathématiques ou dans le domaine visuo-spatial, héritera de la réputation d'infériorité de son groupe dans ces domaines. Nous allons voir combien ce « stigmate » peut être lourd de conséquences.

Comme le fait remarquer Halpern (2000), étudier la relation entre stéréotype et différence de sexe est difficile. En effet, on constate que les différences au niveau cognitif correspondent exactement aux stéréotypes de genre : les femmes auraient de meilleures compétences verbales, et les hommes les surpasseraient dans le domaine spatial et celui des mathématiques. Or les tâches visuo-spatiales, et plus particulièrement la RM, sont justement mieux réussies par les hommes qui déclarent avoir une représentation d'eux-même (*self-concept*) conforme aux traits typiquement masculins (Signorella & Jamison, 1986 ; Hamilton, 1995). Mais les hommes sont-ils encouragés à se perfectionner dans ces domaines du fait que ceux-ci sont associés aux traits de genre masculins — auquel cas les stéréotypes apparaîtraient comme autant de « prophéties auto-réalisatrices » (cf. l'ef-

fet Pygmalion, Rosenthal & Jacobson, 1968) — ou cette association est-elle née du fait que les hommes sont intrinsèquement meilleurs dans le domaine spatial ? Les données de la littérature liant stéréotype et différence de sexe en RM ne permettent hélas pas de trancher, car elles reposent essentiellement sur une approche corrélacionnelle (Halpern, 2000 ; Maccoby & Jacklin, 1974), incapables de conclure quant à la direction de la causalité.

Or l'approche que nous défendons dans cette thèse est précisément qu'il est possible de surmonter cette difficulté méthodologique, et de conclure sans équivoque sur le rôle des facteurs sociaux. En effet, le paradigme expérimental proposé par Steele et Aronson (1995) envisage la manipulation du stéréotype comme une variable expérimentale. Selon cette approche, les stéréotypes de genre peuvent avoir une influence non seulement au cours de l'expérience socialisante des individus, mais également — et surtout — de manière très ponctuelle, très localisée dans le temps, au moment même où l'individu est engagé dans une tâche, notamment de RM. Nous consacrons la section suivante à la présentation détaillée de ce paradigme.

2.2 Interférence, ou « menace », du stéréotype

2.2.1 Une Approche Situationnelle

Steele et Aronson (1995) font l'hypothèse d'un mécanisme d'intervention du stéréotype qui rompt avec l'approche socialisante classique. Selon ces auteurs, toute personne cible d'un stéréotype dans un domaine donné, confrontée à une évaluation dans ce domaine, — par exemple les femmes face à un test d'aptitude spatiale — est susceptible de nourrir la crainte d'un échec qui confirmerait alors, pour eux-mêmes ou aux yeux d'autrui, le stéréotype. Ces craintes pourraient perturber les processus cognitifs mobilisés lors de la résolution du test, et amener justement cette personne ciblée par un stéréotype à échouer.

Nous parlerons alors d'*interférence du stéréotype*, Steele et Aronson ont quant à eux opté pour une désignation plus phénoménologique, évoquant une « menace » du stéréotype (*stereotype threat*), mettant l'accent sur l'enjeu créé par l'activation du stéréotype pour l'identité du sujet.

L'expérience princeps de Steele et Aronson (1995) a démontré la réalité de ce phénomène chez les noirs américains ciblés par un stéréotype leur prêtant une intelligence moindre qu'aux blancs. Confrontés à un test verbal, ostensiblement présenté comme un test de leurs capacités intellectuelles, les noirs montraient une performance inférieure à celles des blancs. Pourtant, lorsque le même test était présenté comme un ensemble de simples problèmes verbaux, noirs et blancs réalisaient la même performance. Une légère variation dans la présentation du test modifiait radicalement pour les participants noirs la signification de la situation d'évaluation. La présentation du test comme diagnostique de l'intelligence les renvoyait à la possibilité de confirmer le stéréotype, créant un risque en effet de nature à interférer avec la performance.

Depuis, ce résultat a maintes fois été répliqué, auprès de divers groupes ciblés par autant de stéréotypes différents. On compte aujourd'hui des centaines d'études sur l'interférence du stéréotype, dont un certain nombre de synthèses (Désert, Croizet, & Leyens, 2002 ; Schmader et coll., 2008 ; Shapiro & Neuberg, 2007 ; Walton & Cohen, 2003 ; Wheeler & Petty, 2001). Le phénomène semble exister pour tous les groupes sociaux, du moment qu'un stéréotype négatif pertinent quant à la situation d'évaluation entre en jeu. Ce phénomène a dans un premier temps été mis en évidence relativement au groupe ethnique de certains sujets. Initialement, Steele et Aronson (1995) ont mis en évidence la survenue du phénomène chez des étudiants afro-américains, mais d'autres groupes souffrant de la même réputation d'infériorité intellectuelle peuvent être touchés, comme les hispaniques et notamment les femmes hispaniques (Gonzales, Blanton, & Williams, 2002), ou encore des africains dans un contexte de simulation d'entretien d'embauche (Klein, Pohl, & Ndagijimana, 2007). D'autres auteurs ont montré les conséquences négatives de l'intervention

de stéréotypes défavorables aux femmes en mathématiques (Spencer et coll., 1999 ; Ben-Zeev, Carrasquillo, et coll., 2005 ; Cadinu et coll., 2005 ; Shih, Pittinsky, & Ambady, 1999), mais aussi dans le domaine visuo-spatial (en relation avec la géométrie, Huguet & Régner, 2007, 2009), ou encore en sport (Chalabaev, Sarrazin, Stone, & Cury, 2008). Les observations les plus frappantes sont celles montrant l'existence de situations où les hommes blancs, en dépit de leur statut social *a priori* dominant, montrent une infériorité lorsque la comparaison sociale leur est défavorable. Ils produisent en effet des performances plus faibles face à des asiatiques en mathématiques (Aronson et coll., 1999), face à des sujets de sexe féminin sur un test de perception émotionnelle (Leyens, Désert, Croizet, & Darcis, 2000), face à des afro-américains (Stone, Lynch, Sjomeling, & Darley, 1999) ou des hispaniques (Stone, 2002) dans le domaine sportif. Enfin des individus réputés moins intelligents de par leur statut économique et social (Croizet & Claire, 1998 ; Croizet et coll., 2004), ou leur parcours éducatif (Croizet & Dutrévis, 2004), s'avèrent également sensibles au stéréotype négatif dont ils font l'objet. En bref, l'influence théorisée par Steele et ses collaborateurs bénéficie aujourd'hui de nombreuses démonstrations expérimentales.

Une démonstration particulièrement pertinente pour notre propos est celle de Spencer et coll. (1999) auprès d'étudiants des deux sexes, tous performants en mathématiques. Des extraits d'un test standard dans ce domaine (SAT-M) leur était proposé selon trois conditions. (1) Dans l'une d'elle les étudiants apprenaient que le test était propice à une différence de sexe (condition *d'activation*). (2) Dans une autre condition ils apprenaient qu'aucune différence de sexe n'était attendue (condition de *falsification*). (3) Enfin, dans une autre condition encore, la question des différences de sexe n'était nullement évoquée (condition *contrôle*). Dans la condition d'activation explicite du stéréotype, Spencer et coll. observaient bien la différence de sexe attendue à l'avantage des hommes, tout comme dans la condition dite de contrôle, exactement comme si dans cette condition apparemment neutre le stéréotype était implicitement activé. Cela d'autant que dans la condition de falsification la différence de sexe était annulée. L'apparition de cette différence même dans la condition où aucune information explicite évoquant le stéréotype n'avait été dé-

livrée, indique certainement que l'évidente nature mathématique du test était suffisante pour activer en mémoire des connaissances en rapport avec le stéréotype de genre, et de là, conduire les femmes à une performance inférieure. Il apparaît ainsi que la référence explicite au stéréotype n'est pas nécessaire pour que s'exprime l'interférence imaginée par Steele et ses collaborateurs. Par ailleurs, il est frappant de constater dès 1999 à quel point la différence sexe en mathématiques apparaît malléable au point d'apparaître ou de disparaître selon la nature des instructions délivrées au sujet au moment du test.

2.2.2 Mécanismes

On peut distinguer trois types d'approche parmi les études cherchant à expliquer l'action du stéréotype interférent. Le premier consiste à chercher les conditions favorables à la survenue du phénomène, le deuxième tente d'établir quelles actions entreprendre pour prévenir sa survenue. Le dernier enfin tente d'identifier les mécanismes cognitifs ou émotionnels explicatifs.

Conditions propices à une interférence du stéréotype

On peut distinguer plusieurs facteurs critiques pour cette interférence. L'un de ces facteurs, bien évidemment, est l'appartenance de l'individu à un groupe ciblé par un stéréotype négatif, dont l'accessibilité en mémoire varie selon les contextes (Steele & Aronson, 1995). Ensuite, l'importance que l'individu attache au domaine de compétence évoqué dans le stéréotype. Pour bien saisir l'importance de ce facteur, il ne faut pas oublier que, selon la théorie, l'effet interférent du stéréotype est avant tout dû au fait que l'éventualité d'un échec menace l'identité du sujet. Or une telle menace ne peut survenir que si les compétences évaluées par le test sont constitutives de cette identité (domaine central pour la définition de soi, cf. Aronson et coll., 1999 ; Leyens et coll., 2000). Mais cela n'est encore pas suffisant : il faut aussi que le sujet se sache performant dans le domaine

de compétence évalué. En effet, des individus d'un niveau médiocre (ou mauvais) face à ce type de test n'ont aucune raison de redouter un échec auquel ils sont, en quelque sorte, habitués. Par contre, l'enjeu devient important pour les sujets les plus performants qui par définition ont beaucoup à perdre en cas d'échec. La frange de la population sensible à l'influence délétère du stéréotype est en effet constituée par les individus les plus compétents dans le domaine évalué¹. Enfin, une dernière condition nécessaire pour que le phénomène d'interférence survienne est que les tests proposés présentent une certaine difficulté, et donc un certain risque d'échec (Ben-Zeev, Carrasquillo, et coll., 2005 ; Spencer et coll., 1999).

À première vue, le fait que les sujets les plus performants soient sensibles à l'effet négatif du stéréotype peut paraître paradoxal. En effet, ces individus en réussite dans le domaine ciblé par le stéréotype sont par définition contre-stéréotypiques. Ce paradoxe s'exprime également dans des résultats plus récents auprès de jeunes collégiennes qui, tout en affichant des croyances contre-stéréotypiques (c.-à-d. supériorités des filles sur les garçons en mathématiques), souffrent elles aussi de l'interférence en question (Huguet & Régner, 2009). Conformément à une proposition clef de Steele (1997), l'adhésion au stéréotype ne semble donc pas nécessaire à cette interférence.

Inductions expérimentales

Les manipulations expérimentales à même de provoquer — ou à l'inverse, de prévenir — le phénomène, sont relativement bien documentées. On désignera par le terme d'*induction* ces manipulations qui visent à contrôler l'activation du stéréotype. Ces inductions peuvent revêtir différentes formes, de l'évocation explicite du stéréotype dans les instructions jusqu'à son activation implicite. Par exemple le stéréotype peut être implici-

1. Ainsi, par exemple, Steele et Aronson (1995) testent des afro-américains étudiants de l'université Stanford avec de hauts scores aux tests standardisés d'entrée à l'université (SAT), et Spencer et coll. (1999) avaient sélectionné des étudiantes dont le score à des tests standardisés de mathématiques (SAT-M ou ACT-M) était supérieur au 85ème percentile de la distribution de performance.

tement activé par le biais de la présence majoritaire de membres de l'exo-groupe, qui suffit à rendre saillante l'appartenance à un groupe ciblé par le stéréotype. Ainsi Sekaquaptewa et Thompson (2003) ont montré que des femmes qui passaient seules un test de mathématiques parmi plusieurs hommes étaient moins performantes que celles confrontées au test parmi d'autres femmes. On se souvient aussi que, dans l'expérience de Steele et Aronson (1995), le stéréotype était également activé implicitement par le caractère diagnostique du test. En outre, les résultats de Spencer et coll. (1999) évoqués précédemment montrent que certains tests peuvent apparaître très clairement évaluatifs d'une compétence donnée. Par exemple un test de mathématiques est directement perçu comme une évaluation des compétences en mathématiques, surtout s'il est difficile. De même un test de rotation mentale 3D peut difficilement passer pour autre chose qu'un test des compétences visuo-spatiales. Dans de tels cas, il est pratiquement impossible de manipuler, comme Steele et Aronson (1995) l'avaient fait, la diagnosticité du test. Le plan expérimental proposé par Spencer et coll. (1999), qui proposent de comparer une condition de *falsification* aux conditions d'*activation* (explicite ou implicite) du stéréotype, est alors pour nous d'un intérêt capital.

D'autres auteurs ont envisagé des manipulations alternatives à la falsification pour contrer l'effet délétère du stéréotype. La recherche d'interventions permettant d'annuler l'effet d'interférence du stéréotype est importante non seulement parce que la demande sociale est forte concernant de telles interventions, mais leur mise en évidence permet d'affiner la compréhension même des processus en jeu (Aronson, Cohen, McColskey, & Regional Educational Laboratory Southeast, 2009). Quelques pistes explorées sont le rôle de l'affirmation de soi (G. L. Cohen, Garcia, Apfel, & Master, 2006 ; Martens et coll., 2006 ; Schimel, Arndt, Banko, & Cook, 2004), celui de l'activation de dimensions de l'identité non impliquées par le stéréotype (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004 ; Shih et coll., 1999 ; Shih, Bonam, Sanchez, & Peck, 2007) ou encore l'effet protecteur des attributions externes de type auto-handicap, etc. (Ben-Zeev, Fein, & Inzlicht, 2005 ; Johns, Inzlicht, & Schmader, 2008 ; Keller, 2002). En dépit de l'importance

de ces travaux, nous nous en tiendrons à la proposition de falsification de Spencer et coll. (1999), préférant nous concentrer sur les mécanismes cognitifs sous-jacents à l'interférence du stéréotype plutôt que d'étudier les méthodes d'intervention pour elles-mêmes.

Mécanismes de l'interférence du stéréotype

Bien que le *stereotype threat* soit à l'origine d'un nombre très important d'études, les processus sous-jacents ne sont toujours pas connus avec précision (Beilock, Rydell, & McConnell, 2007 ; Mayer & Hanges, 2003 ; Schmader et coll., 2008 ; Wheeler & Petty, 2001). L'axe de recherche qui a concentré la majorité des efforts a consisté à identifier les mécanismes cognitif et / ou émotionnels sous-jacents à l'interférence en question.

Steele et Aronson (1995) et Steele (1997) décrivent le phénomène en référence à une menace, une crainte. L'anxiété a donc été un médiateur fréquemment étudié, mais avec des résultats peu concluants, sinon contradictoires (Aronson et coll., 1999 ; Mayer & Hanges, 2003 ; Osborne, 2001 ; Spencer et coll., 1999 ; Stone et coll., 1999). Schmader et Johns (2003) ont quant à eux situé le mécanisme d'interférence au niveau de la capacité de la mémoire de travail qui apparaît affaiblie dans les conditions propices à l'activation du stéréotype. La voie était ainsi ouverte à une approche des mécanismes en termes cognitifs plutôt qu'émotionnels. Peu après, Croizet et coll. (2004) ont montré à l'aide d'observables physiologiques (variations du rythme cardiaque) que les effets d'interférence s'accompagnaient d'une augmentation de la charge mentale.

Ces observations restent cohérentes avec l'hypothèse initiale de Steele et Aronson (1995) : si la quantité de ressources cognitives disponibles est entamée par l'activation du stéréotype, c'est parce que les craintes expérimentées par les sujets les conduisent à nourrir des pensées négatives, des ruminations, qui, s'accumulant, consomment les ressources cognitives. Si Keller et Dauenheimer (2003) on pu conclure que l'interférence en question était bien liée à l'émergence d'émotions et de pensées négatives, Cadinu et coll. (2005) ont

montré que des femmes confrontées à un test difficile de mathématiques nourrissaient de plus en plus de pensées négatives au fur et à mesure qu'elles progressaient dans le test, le nombre de pensées négatives rapportées médiatisant l'effet du stéréotype. Il semble donc bien que, dans cette situation, les ressources cognitives sont entamées par des pensées négatives. Il n'est pas exclu, comme le proposent d'ailleurs Schmader et coll. (2008), que ce phénomène d'intrusion de pensées négatives s'exprime en priorité voire même exclusivement chez les sujets dont la capacité de contrôle cognitif est relativement faible. Or cette capacité recouvre celle de la mémoire de travail (Ilkowska, Engle, & Hoyle, 2010 ; Kane et coll., 2007). D'où la possibilité d'un effet du stéréotype plus marqué chez les sujets à faible empan en mémoire de travail, relativement à ceux à plus fort empan. C'est en effet exactement ce qu'indiquent les résultats récents de Régner et coll. (in press) auprès d'étudiant(e)s ingénieurs soumis au test de raisonnement des Matrices de Raven (cf. aussi Croizet et coll., 2004, pour une illustration expérimentale de la menace du stéréotype avec de telles Matrices). Cet ensemble de travaux offre un cadre qui nous permet d'envisager l'intervention du stéréotype dans des activités moins intégrées que celle liée à la résolution de problèmes mathématiques.

2.2.3 Différence de sexe visuo-spatiale et stéréotype

Jusqu'à récemment, la relation entre différence de sexe et intervention du stéréotype a surtout été étudiée en mathématiques plutôt qu'au niveau des aptitudes visuo-spatiales. Les mathématiques ont peut-être plus attiré l'attention car la demande sociale — du fait de leur rôle crucial dans la sélection des étudiants pour l'accès aux filières d'enseignements scientifiques et techniques — est très forte (Aronson et coll., 2009). En outre, par rapport à la rotation mentale, les mathématiques constituent un domaine cognitivement plus intégré, donc peut-être perçu comme plus perméable à l'intervention de croyances et représentations tels que les stéréotypes de sexe. Robuste, portant sur un processus cognitif d'assez bas niveau, la différence de sexe en RM a pu paraître hors de portée d'une telle

influence, d'autant qu'un stéréotype de sexe au niveau spatial semble moins évident, bien que son existence soit manifeste (cf. la fameuse incapacité des femmes à lire une carte routière), ou encore le fait qu'on attende moins d'une femme qu'elle excelle dans le domaine spatial (Alington et coll., 1992 ; Blakemore et coll., 2009 ; Hamilton, 1995 ; Scali & Brownlow, 2000 ; Signorella et coll., 1989).

Pourtant de nombreux éléments suggèrent que la différence de sexe observée en RM peut être, au moins pour partie, due à un phénomène de type interférence du stéréotype :

1. nous venons de le mentionner, il existe un stéréotype invoquant la relative incompetence des femmes dans le domaine visuo-spatiales ;
2. la différence de sexe s'observe essentiellement sur les tests de RM les plus difficiles (Collins & Kimura, 1997 ; Kimura, 1999), notamment en 3D (Voyer et coll., 1995), or la difficulté du test est un élément essentiel pour l'occurrence du *stereotype threat* ;
3. la performance des femmes en RM montre une certaine variabilité (Quaiser-Pohl & Lehmann, 2002), qui suggère l'intervention d'un modérateur non-identifié. De par sa versatilité, l'effet d'interférence du stéréotype pourrait jouer ce rôle ;
4. la différence de sexe en RM est d'autant plus importante que l'on compare des hommes et des femmes d'un bon niveau (Casey, Nuttall, et coll., 1995), or les effets de l'intervention du stéréotype sont particulièrement forts chez les individus les plus capables (identifiés au domaine) ;
5. la différence de sexe en RM est caractéristique des différences localisées au niveau des tâches qui nécessitent un traitement actif en mémoire de travail (Cornoldi & Vecchi, 2003 ; Halpern & Wright, 1996), or l'effet de l'interférence du stéréotype semble surtout localisé en mémoire de travail, entamant les ressources disponibles (Schmader et coll., 2008).

Le paradigme de *stereotype threat* apporte donc un éclairage cohérent avec de nombreux aspects de la différence de sexe en RM. Ce n'est pourtant que très récemment qu'ont été publiées des études confrontant intervention du stéréotype et RM 3D².

2.3 Stéréotype et rotation mentale

2.3.1 Sharps, Welton et Price (1993) : des précurseurs

Facteurs non-cognitifs et rotation mentale

Plusieurs études ont tenté de mettre à jour la relation entre stéréotype et différence de sexe en RM. Sharps, Welton, et Price (1993) furent résolument des précurseurs dans ce domaine. Ignorant la théorie du *stereotype threat* — et pour cause, elle ne sera formulée que deux ans plus tard par Steele et Aronson, en 1995 — Sharps et coll. proposeront néanmoins un paradigme très proche de celui qui sera mis en œuvre par les découvreurs du phénomène d'interférence. Partant du constat que la différence de sexe s'exprimait différemment en fonction du type de tâche, ces auteurs ont essayé de déterminer si cette variation était imputable à des facteurs cognitifs exclusivement — des tâches différentes mobilisant des processus cognitifs différents — ou si des facteurs décrits comme « non-cognitifs », d'ordre socio-psychologique, pouvaient aussi jouer un rôle.

Ces facteurs seraient essentiellement des représentations *construites* par le sujet quant à sa capacité à réussir une tâche donnée, en fonction de son expérience passée, de ses croyances, des croyances partagées dans son environnement social, des stéréotypes, etc. Sharps et coll. ont manipulé l'habillage du test via les instructions pour contrôler ces

2. Quelques auteurs ont réalisé très récemment des études avec des tests comme la *figure de Rey* (Huguet & Régner, 2007, 2009) ou le *paper folding test* (Keller & Sekaquaptewa, 2008). Ces tests ne montrant généralement qu'une différence de sexe de petite taille lorsqu'ils montrent une différence, nous ne les avons pas inclus dans cette revue où nous nous sommes restreints aux tests de RM en 3D.

facteurs en accord avec les critères classiques de la méthodologie expérimentale, et pour éviter les écueils d'une approche corrélationnelle.

Pour Sharps et coll., du fait de leur réputation d'infériorité dans le domaine spatial, les femmes devraient se montrer moins performantes à un test de RM lorsqu'il est présenté d'une manière qui souligne son caractère spatial, par rapport à une situation où cette dimension spatiale est atténuée. C'est exactement ce qu'ont vérifié Sharps et coll.. Dans leur expérience, des étudiants des deux sexes étaient confrontés à une section du MRT de Vandenberg et Kuse (1978). À la moitié des sujets, le test était présenté comme « une évaluation de vos capacités spatiales, la capacité à raisonner et à résoudre des problèmes impliquant des objets matériels dans l'espace. »³ (condition spatiale). À l'autre moitié il était dit que ce test permettait « une évaluation de vos capacités intellectuelles, votre capacité à raisonner et à résoudre des problèmes. »⁴ (condition non-spatiale). La différence de sexe était significative dans la condition spatiale, alors qu'elle ne l'était plus dans la condition non-spatiale.

Cette étude fut la première à révéler la sensibilité d'une différence de sexe (jugée par ailleurs très robuste) à l'habillage du test, et donc en l'occurrence aux significations attribuées à l'activité cible par les sujets en fonction de leur groupe d'appartenance. Si les auteurs avouaient ne pas savoir quel mécanisme proposer pour expliquer ce phénomène, ils voyaient néanmoins dans leurs résultats la preuve de la perméabilité des fonctions cognitives au contexte, notamment social, dans laquelle cette cognition prend place, évoquant une approche de type *cognition située*. En 1994, Sharps et coll. reproduiront ces résultats et préciseront deux points essentiels : l'importance du rôle de la difficulté de la tâche pour la survenue de l'effet, et l'implication spécifique des stéréotypes de sexe.

3. *The following is an evaluation of your spatial abilities, of your abilities to reason and solve problems regarding physical objects in space.*(Sharps et coll., 1993)

4. *The following is an evaluation of your mental abilities, of your abilities to reason and solve problems.*(Sharps et coll., 1993)

Stéréotype, difficulté et différence de sexe

Pour établir comment difficulté et facteurs non-cognitifs interagissaient, Sharps et coll. ont proposé à différents groupes de participants un test de RM proposant des problèmes de difficulté variable, selon la nature des items constituant les problèmes : des figures simples en deux dimensions, des figures en deux dimensions rappelant celles utilisées dans le test de Vandenberg et Kuse (1978), des figures familières en 3D (cheval, brosse...) et enfin des figures 3D abstraites tirées du MRT de Vandenberg et Kuse (1978). Les tests étaient proposés à des hommes et des femmes selon les modalités vues précédemment : présentation *spatiale* ou *non-spatiale*. La différence de sexe ne fut observable que pour les items abstraits en 3D et en 2D (mais extrêmement ténue et avec de très hauts scores pour les deux sexes dans le second cas). En fait, la différence la plus importante apparaissait sur les problèmes abstraits en 3D avec les rotations les plus importantes (difficiles). Parallèlement, l'interaction entre le sexe des participants et le type de présentation du test n'était significative que pour les items en 3D de type Shepard et Metzler, ceux-là mêmes qui permettent classiquement de mettre à jour le plus clairement la différence de sexe en RM.

Pour s'assurer que l'élément contextuel à l'origine de l'interaction était bien lié au stéréotype de sexe, les auteurs ont dans une seconde étude modifié la présentation du test, manipulant la diagnosticité non plus selon un axe spatial / non-spatial, mais un axe masculin / féminin. Le test était soit présenté comme un outil permettant de déceler des capacités critiques pour réussir dans des activités (stéréo)typiquement masculines : ingénierie aérienne, nucléaire, armement, pilotage, etc., soit comme un révélateur des capacités liées à des activités typiquement féminines : architecture intérieure, décoration, confection, stylisme, etc. Le test était bien entendu le même dans les deux conditions : une version standard du MRT de Vandenberg et Kuse, pour laquelle une interaction entre le sexe des participants et le type d'instruction s'est une fois encore montrée significative. Cependant, cette interaction était due au fait que les hommes dans la condition « masculine »

affichaient un niveau de performance supérieur à toutes les autres conditions. La manipulation expérimentale n'avait aucune incidence pour les femmes. Seule la performance des hommes était modulée.

Menace du stéréotype et rotation mentale : un rendez-vous manqué

Il est troublant de constater que les résultats assez forts et originaux établis par Sharps et coll. (1993) ont été, dans les années qui ont suivi leur publication, peu cités dans la littérature sur les différences cognitives entre les sexes. Cela est d'autant plus regrettable lorsque l'on constate à quel point leur paradigme est proche de celui mis en œuvre par Steele et Aronson deux ans plus tard dans leur étude princeps de 1995. L'approche et les résultats de Sharps et coll. s'organisent autour d'éléments qui sont centraux dans la théorie du *stereotype threat* : le rôle du contexte notamment dans sa composante sociale, la possibilité de manipuler expérimentalement ce contexte par le biais de la signification attribuée à la tâche, l'interaction entre cette signification et l'identité du sujet, le fait que cette interaction ne soit perceptible qu'au niveau de la performance dans les tâches les plus difficiles. Si la théorie de Steele et Aronson formulée quelques années plus tard a pu constituer un cadre tout indiqué pour intégrer les différentes propositions (pour ne pas dire intuitions) et résultats de Sharps et coll., il semblerait que le rendez-vous a été manqué, certainement du fait de la discrétion de la publication de leurs résultats d'une part, et d'un certain cloisonnement entre deux domaines différents mais néanmoins complémentaires de la psychologie sociale et de la psychologie cognitive. C'est ce rapprochement que nous proposons dans cette thèse.

Il faudra attendre 2006, soit une décennie après la première formulation de la théorie du *stereotype threat* et la publication des résultats de Sharps et coll. (1993, 1994), pour qu'une série d'études proposées par différents auteurs envisage la possibilité que ce phénomène puisse expliquer, au moins en partie, la différence de sexe en RM. Jusque là, la

relation entre différence de sexe et interférence du stéréotype avait surtout été étudiée en mathématiques.

2.3.2 Intervention du stéréotype en rotation mentale

Le problème des inductions

Dans le but d'établir une intervention permettant de contrecarrer l'effet délétère du stéréotype, Martens et coll. (2006) ont proposé un test de RM (test de Vandenberg & Kuse, 1978, dans sa version réactualisée par Peters et coll., 1995) à des étudiants des deux sexes (classe entière). Le stéréotype était activé au début du test de RM, présenté comme typiquement mieux réussi par les hommes. Le test était soit proposé directement, soit précédé d'une tâche poussant les participants à s'auto-affirmer (en évoquant plusieurs de leurs qualités). Dans cette condition les femmes obtenaient des scores équivalents à ceux des hommes, alors que dans l'autre condition la différence de sexe en défaveur des femmes était significative. Cependant ces résultats posent deux problèmes d'importance à la fois en regard de la différence de sexe en RM et de l'intervention du stéréotype.

Premièrement les auteurs prétendaient activer le stéréotype en indiquant clairement qu'il existait un stéréotype spécifiant que les femmes réussissaient moins bien le test⁵. Or, dans ces conditions, plus qu'une activation du stéréotype, il s'agit plutôt d'une présentation des résultats attendus. Il y a en effet une différence entre activer une connaissance de l'ordre d'un stéréotype susceptible de guider les attentes des sujets et partant, leur performances, et leur imposer explicitement une information défavorable au groupe d'appartenance. On peut donc douter que ces auteurs se soient placés dans une situation d'intervention du stéréotype — est-ce le stéréotype qui intervient ou l'autorité de l'expérimentateur ?

5. [...]one thing we will look at is how men and women differ in their performance on the test, and how true the stereotype is, or the generally held belief is, that women have more trouble with spatial rotation tasks. (Martens et coll., 2006)

Deuxièmement, l'absence d'une condition de contrôle (sans auto-affirmation et sans feedback), qui aurait permis de situer les niveaux de performance obtenus dans cette expérience relativement à une condition plus standard où seules les instructions du test sont présentées.

Dans leur étude, Wraga et coll. (2006) ont montré l'effet d'une affirmation arbitraire et normative de la supériorité des femmes en RM⁶ sur la performance des deux sexes à un test de RM. Le test choisi par Wraga et coll. était dérivé du MRT. Les participants devaient indiquer si un point particulier sur un objet semblable à ceux créés par Shepard et Metzler (1971) était toujours visible depuis un autre point de vue. Les sujets devaient donc imaginer la scène comme si c'était eux qui se déplaçaient (*self-rotation*) et non l'objet comme dans le MRT. La différence de sexe en condition « standard » est établie pour cette tâche, les femmes faisant plus d'erreur et répondant plus lentement. L'effet de l'induction, qui ne permet pas de conclure sur l'intervention d'un stéréotype pour les raisons déjà exposées, conduit à une réduction de la différence de sexe, mais pas dans le sens attendu comme nous allons le voir.

Annulation de la différence de sexe : baisse de la performance des hommes

Les résultats obtenus par Wraga et coll. sont représentatifs de la plupart des études qui ont cherché à établir un lien entre interférence du stéréotype et RM. Dans la condition contre-stéréotypique (lorsqu'il est dit aux participants que les femmes sont supérieures aux hommes), les femmes montrent de meilleures performances que dans la condition contrôle (où rien n'est dit au sujet des groupes de sexe) et les hommes des performances plus faibles. Cependant, dans aucune condition — particulièrement dans la condition stéréotypique — les femmes n'atteignent le niveau des hommes dans la condition contrôle. Et si la différence de sexe s'annule dans la condition contre-stéréotypique c'est en raison de la baisse de la

6. *Our previous research has shown that female students consistently outperform male students in this task : They usually respond faster and make fewer errors.* (Wraga et coll., 2006)

performance des hommes. Si cela montre que l'intervention d'une croyance peut altérer une performance à un test de RM, cela ne permet pas d'imputer l'origine de la différence de sexe classiquement observée à un stéréotype, car jamais on ne peut observer que les femmes ont la capacité d'atteindre le meilleur niveau de performance des hommes. En d'autres termes ces résultats ne permettent aucunement de remettre en cause la supériorité des hommes en RM.

On peut mettre en cause le choix de l'induction, et adresser à Wraga et coll. la même critique qu'à Martens et coll. (2006). De plus, peut-être l'évocation d'une différence de sexe — fût-elle en faveur des femmes — active-t-elle justement le stéréotype spatial qui repose sur des années d'expérience sociale et qui serait donc beaucoup plus prégnant pour les femmes que l'induction artificielle proposée par les auteurs, et en modèrerait la portée ?

Dans une étude parue la même année Moè et Pazzaglia (2006) ont opté eux aussi pour des inductions présentant la même faiblesse, mais ont proposé néanmoins une procédure intéressante par rapport à la dynamique de la tâche. Le MRT étant constitué de deux parties de même longueur et d'égale difficulté, ces auteurs ont procédé à une induction entre les deux parties de la tâche, afin de pouvoir comparer l'évolution des performances d'un même individu avant et après l'activation (ou non) du stéréotype. Trois types d'induction ont été pris en considération : le test était présenté soit comme démontrant scientifiquement l'infériorité des femmes dans le domaine spatial⁷ (condition stéréotypique), soit comme démontrant au contraire l'infériorité des hommes⁸ (condition contre-stéréotypique), ou enfin comme un simple test spatial⁹ (condition standard).

On peut relever ici que la condition désignée comme « standard » par Moè et Pazzaglia correspond à la condition sensée déclencher l'intervention du stéréotype pour Sharps

7. *This test measures spatial abilities. Research showed that men perform better than women in this test, probably for genetic reasons. This means that women score lower than men.* (Moè & Pazzaglia, 2006)

8. *This test measures spatial abilities. Research showed that women perform better than men in this test, probably for genetic reasons. This means that men score lower than women.* (Moè & Pazzaglia, 2006)

9. *This test measures spatial abilities. Research showed that spatial ability is very important in everyday life, e.g. to find a route or describe a pathway to someone. This test has been used in the USA and over the last few years also in Europe, in particular Italy, confirming the results obtained with American samples.*

et coll. (1993). Nous constatons en effet que l'induction — manipulation critique lorsqu'il s'agit de mettre en évidence un effet du stéréotype — est envisagée de manière très différente selon les auteurs et, hélas, souvent inadaptée quant à cet objectif. Ainsi, les inductions (stéréotypiques et contre-stéréotypiques) proposées par Moè et Pazzaglia présentent la même faiblesse que celle de Martens et coll. (2006) : en indiquant explicitement et de manière très normative (résultat scientifiquement établi) la direction de la différence de sexe, les auteurs ne font plus intervenir un stéréotype, mais peut-être un facteur contextuel plus général. Les participants sont-ils encouragés ou découragés par l'affirmation de l'expérimentateur ? La situation devient confuse, sinon équivoque, il est alors difficile de tirer des conclusions claires quant aux mécanismes à l'origine de la baisse de performance.

Toujours est-il que Moè et Pazzaglia ont observé une variation significative des performances à la suite des différentes inductions. La performance des femmes a augmenté après l'induction contre-stéréotypique, et celle des hommes après l'induction stéréotypique. Cependant cette variation fut beaucoup plus importante chez les hommes, et si l'on compare les performances des deux sexes (testés dans deux études différentes), on constate des résultats assez similaires à ceux obtenus par Sharps et coll. (1994) et Wraga et coll. (2006) : l'impact des manipulations est surtout localisé au niveau des hommes qui réalisent leur meilleure performance dans la condition qui leur est la plus favorable en regard du stéréotype. Dans ce cas, la réduction — voire l'annulation, mais ce n'est pas clair pour Moè et Pazzaglia qui n'ont pas comparé directement les performances entre les sexes — de l'écart entre hommes et femmes est surtout dû à une baisse de la performance des hommes.

Inductions rigoureuses, résultats mitigés

M. S. McGlone et Aronson (2006) ont approché le problème en agissant très rigoureusement au niveau des inductions. Plutôt que de manipuler la présentation de la tâche, inspirés par Shih et coll. (1999) et leur méthode d'activation des identités sociales mul-

tiples, ils ont préféré activer différentes dimensions de l'identité sociale des participants. À l'aide d'un questionnaire proposé avant le MRT, M. S. McGlone et Aronson portaient l'attention des participants soit sur leur identité de sexe soit sur leur identité d'étudiant(e)s dans une université privée ou enfin sur leur identité de résident du nord des États-Unis.

Conformément à leur hypothèse, M. S. McGlone et Aronson observent bien une interaction entre le type d'identité sociale activé et le sexe des participants, la performance des femmes au MRT étant affectée différemment selon l'identité activée. Les femmes réalisent de plus faibles performances dans la condition où leur identité de sexe était rendue saillante et leurs meilleures performances dans la condition où leur identité d'étudiante était rendue saillante. Cependant il n'y a pas d'écart significatif entre la performance des femmes en condition contrôle et les deux autres conditions. Il n'y a pas non plus de variation significative de la performance des hommes à travers les conditions, et cette performance est toujours supérieure à celle des femmes.

En bref, l'étude de M. S. McGlone et Aronson, en dépit de ses qualités intrinsèques ne permet pas pour autant d'expliquer la différence de sexe en RM par l'intervention d'un stéréotype. Certes l'expression de la performance des femmes apparaît sensible aux manipulations du contexte social, mais pas dans des proportions suffisantes pour faire la différence avec un contexte plus « neutre ».

En conclusion, toutes les études sur l'interférence du stéréotype et la RM présentées jusqu'ici montrent unanimement que la performance peut être affectée par des facteurs contextuels d'origine sociale, dont le lien avec le traitement cognitif de l'information visuo-spatiale n'est pas intuitif. Cependant, les résultats sont difficilement comparables d'une étude à l'autre tant les manipulations du contextes (inductions) sont variées, et même contradictoires quelquefois : certains auteurs (Sharps et coll., 1993) choisissent de présenter le MRT comme un test « spatial » pour activer le stéréotype alors que cette présentation est associée à une condition contrôle par d'autres (Moè & Pazzaglia, 2006). Dans certains études les inductions altèrent la performance des hommes, et si la différence de

sexe est annulée, c'est parce que ceux-ci voit leur performance baisser au niveau de celle des femmes (Moè & Pazzaglia, 2006 ; Sharps et coll., 1994). Pour d'autres c'est la performance des femmes qui varie selon le type d'induction, mais alors la comparaison avec la performance des hommes n'est pas toujours très claire (M. S. McGlone & Aronson, 2006), ne permettant pas d'isoler une intervention de nature à annuler la différence de sexe.

Stéréotype et observables biologiques

Nous l'avons vu au chapitre précédent, les études portant sur la RM sont nombreuses à avoir accordé de l'importance aux facteurs biologiques. Face à l'apport de la théorie de l'interférence du stéréotype et ses résultats, certains chercheurs ont tenté de recueillir des observables de nature biologique tout en manipulant le contexte de la tâche. L'idée étant, dans le cadre d'une approche psycho-bio-sociale (Halpern, 2000), que les variations de performance observées et provoquées expérimentalement sont dues à une interaction entre les composantes sociales de la situation, l'identité de l'individu et ses caractéristiques innées, biologiquement déterminées. Une autre interrogation émerge également : à quel niveau le contexte social, le stéréotype, intervient-il ? Perturbe-t-il directement les processus cognitifs eux-mêmes (ressources disponibles, c.f. Schmader & Johns, 2003) ou s'agit-il plutôt d'une altération plus générale de l'état de l'individu (stress, anxiété ...) ? Des observables biologiques se révéleraient alors particulièrement précieuses pour apporter un éclairage précis sur l'évolution de l'état de l'individu au cours de la tâche.

Stéréotype et IRMf Wraga et coll. (2007) ont utilisé la technique d'IRMf afin de vérifier quelles étaient les régions cérébrales activées lors d'une tâche de RM suite à différentes inductions. Les zones activées indiquaient-elles une plus grande charge émotionnelle et / ou une plus grande charge mentale ? Les auteurs ont proposé à de jeunes femmes la tâche (*self-rotation*) et les inductions validées dans leur étude précédente (Wraga et coll., 2006, c.f. p. 68), qui avait montré la possibilité d'une modulation de la performance des

femmes en RM. Wraga et coll. rapportent encore une fois une variation de la performance en fonction des conditions conforme aux hypothèses.

Ces auteurs ont également pu établir que les zones cérébrales activées au cours de la tâche étaient différentes selon le contexte dans lequel était placé le sujet. En comparant les activations cérébrales dans la condition « stéréotypique » à celles observées en condition contrôle, il est apparu que les zones particulièrement activées correspondaient au traitement de l'information sociale et des émotions en rapport à soi (gyrus orbital droit) et aux affects (partie ventrale droite du cortex cingulaire antérieur). Les auteurs rapportent par ailleurs une corrélation positive entre l'activation de ces zones et le nombre d'erreurs en RM. La comparaison entre la condition contrôle et la condition « contre-stéréotypique » indiquait une mobilisation plus importante de la partie ventrale du cortex préfrontal, souvent reliée à une mobilisation accrue des ressources en mémoire de travail.

Du fait que des participants masculins n'ont pas été testés, il faut être prudent dans les conclusions que cette étude permettrait de tirer quant à la différence de sexe en RM. Cependant, et malgré toutes les réserves sur les choix des inductions par Wraga et coll., celles-ci ont eu un impact indéniable sur l'activité des sujets et leur expérience au cours de la tâche : l'exposition à une information négative sur le groupe d'appartenance (infériorité des femmes) a semble-t-il centré les participantes sur cette information sociale et leurs propres émotions, perturbant la réalisation de la tâche. À l'inverse, une croyance positive leur aurait permis de se détacher de cette préoccupation et de mobiliser plus de ressources pour mener la tâche à bien. Ce sont des résultats extrêmement encourageants qui montrent qu'il faut se garder d'une vision trop simpliste des relations entre les processus cognitifs et leur substrat biologique. Comme nous l'avons vu au chapitre précédent, de nombreuses données dans la littérature (Gizewski, Krause, Wanke, Forsting, & Senf, 2006 ; Jordan et coll., 2002 ; Walter, Roberts, & Brownlow, 2000 ; Weiss et coll., 2003) indiquent que la différence de sexe semble médiatisée par des activations cérébrales différentes, suggérant une organisation anatomo-fonctionnelle distincte selon les sexes, certainement génétiquement

acquise. Les données de Wraga et coll. obligent à revoir cette approche. Si la RM est bien une activité cognitive fondamentale dont la réalisation semble corrélée avec l'activation de zones cérébrales précises, il ne faut pas oublier que le cerveau fonctionne de manière massivement distribuée. Ainsi, le traitement d'informations a priori complètement détachées de l'information visuo-spatiale peut, en parallèle, mobiliser d'autres circuits cérébraux — notamment ceux mobilisés par les processus émotionnels et d'origine sociale —, perturbant la réalisation de la RM. Ainsi l'activité cérébrale, à l'image des processus purement psychologiques, doit être envisagée non pas comme figée, mais hautement perméable au contexte, notamment social, et à l'historicité du sujet.

2.4 Questions non résolues, nos objectifs

Nous avons détaillé plusieurs études qui ont tenté de mettre en évidence une intervention du stéréotype responsable de la différence de sexe observée en RM. Les résultats sont encourageants, ils montrent effectivement que la manipulation du stéréotype permet de moduler la performances des hommes et des femmes à un test de RM, voire d'annuler la différence de sexe dans certaines études (Hausmann et coll., 2009 ; Martens et coll., 2006 ; Sharps et coll., 1994 ; Wraga et coll., 2006). Malheureusement, ces dernières ne montrent pas que les femmes peuvent atteindre le niveau de performance des hommes, ce qui prouverait qu'elles ont les mêmes capacités en RM — simplement entravées quand le stéréotype défavorable est activé —, mais seulement que les hommes, dans les conditions où le stéréotype ne leur est pas favorable, réalisent les mêmes performances que les femmes. En d'autres termes, ces résultats indiquent pourquoi les hommes peuvent être meilleurs sous l'effet d'un stéréotype qui leur est favorable (*stereotype lift*, cf. Walton & Cohen, 2003), mais pas pourquoi les femmes leur sont typiquement inférieures (*stereotype threat*).

La logique qui sous-tend le paradigme de mise en évidence de l'interférence du stéréotype est la suivante : il s'agit de construire des situations où le stéréotype est activé, que l'on comparera à des situations où il est en principe annulé (falsification), et à des situations où aucune information sociale n'est donnée au sujet (Spencer et coll., 1999 ; Steele & Aronson, 1995). *A priori*, les études qui ont montré la différence de sexe classique en RM n'étaient pas conçues pour placer le sujet dans une situation où le stéréotype aurait été clairement activé, mais plutôt, en regard des standards de la méthodologie expérimentale, dans des conditions sensément neutres. En fait, selon la théorie de l'interférence du stéréotype, cette condition « neutre » ne l'est pas vraiment car la nature évidemment spatiale de la tâche peut être suffisante pour activer implicitement le stéréotype, qui peut alors expliquer la différence de sexe observée. Ainsi, prouver l'intervention d'un stéréotype dans de telles conditions doit se faire en deux temps. D'abord prouver que ces conditions soit-disant neutres sont équivalentes à une situation où le stéréotype est clairement activé. Puis il faut prouver aussi qu'il existe une situation où l'annulation du stéréotype conduit les femmes à réaliser d'aussi bonnes performances que les meilleures performances des hommes.

À de rares exceptions près (M. S. McGlone & Aronson, 2006 ; Hausmann et coll., 2009), les études présentées jusqu'ici ne remplissent pas ces critères. Pratiquement aucune étude ne propose une condition contrôle qui reproduise les conditions standards de passation du test de RM, ce qui rend la comparaison à un « niveau de base », et au reste de la littérature, confuse. D'autre part, nous critiquons vigoureusement les inductions majoritairement choisies (Martens et coll., 2006 ; Moè & Pazzaglia, 2006 ; Wraga et coll., 2006, 2007) car selon nous elles ne permettent pas de s'assurer qu'elles ont bien activé le stéréotype visé, mais peut-être d'autres croyances. Pourtant Spencer et coll. (1999), confrontés à des tests de mathématiques qui par leur évidente nature mathématique étaient à même d'activer implicitement un stéréotype négatif, ont proposé une procédure simple, rigoureuse et efficace de falsification du stéréotype. Cette procédure peut aisément être mise en œuvre pour des tests de RM dont la nature spatiale est évidente.

Nous proposons donc dans cette thèse un ensemble d'études conçues pour répondre à ces faiblesses et clarifier le rôle d'une interférence du stéréotype dans la différence de sexe observée en RM. Nous proposons dans un premier temps deux études comportementales mettant en œuvre le classique MRT de Vandenberg et Kuse (1978) auprès de deux populations différentes : collégiens (chapitre 3, p. 77) et étudiantes (chapitre 4, p. 116). Ces études auront pour but de valider, auprès d'effectifs importants, les inductions et la mise en place d'un plan expérimental comprenant une condition contrôle « standard » au regard de la littérature en RM, qui sera contrastée avec une condition où le stéréotype sera explicitement activé et une condition où le stéréotype sera falsifié.

Notre objectif est également de mettre à jour les mécanismes cognitifs qui pourraient expliquer, au moins en partie, l'action du stéréotype sur les performances en RM. Nous avons vu que les ressources attentionnelles étaient jugées critiques : les pensées négatives suscitées par l'activation du stéréotype semblent capturer une part suffisante de ces ressources pour provoquer une baisse de performance. Nous proposons alors une étude présentant une version informatisée d'un test de RM qui, avec notamment la possibilité de recueillir et donc d'analyser les temps de réponses, nous permettrons d'inférer quels sont les processus (prise de décision, attention, etc.) affectés par l'activation du stéréotype (chapitre 5, p. 143).

Chapitre 3

Stéréotype et rotation mentale 3D, une approche expérimentale

3.1 Objectifs de l'étude 1

Le but de l'étude présentée dans ce premier chapitre expérimental est de montrer la pertinence du paradigme de ST dans le domaine des différences de sexe au niveau des aptitudes visuo-spatiales, notamment en RM. En effet, nous avons vu précédemment que les résultats des études confrontant différence de sexe dans le domaine visuo-spatial (RM) et interférence du stéréotype sont relativement versatiles. Il convient donc, dans un premier temps, de s'assurer qu'il est possible de construire avec précision un contexte expérimental permettant de manipuler les conditions d'activation du stéréotype qui cible les femmes dans le domaine visuo-spatial. La falsification explicite du stéréotype peut-elle réduire, sinon annuler, la forte différence de sexe systématiquement observée jusqu'à présent en RM ? Les observations de Spencer et coll. (1999) faisant état d'une disparition de la différence de sexe dans le domaine des mathématiques — activité cognitive nettement

plus intégrée que la RM — sont-elles toujours valides pour une activité de traitement de l'information plus élémentaire, de plus bas niveau, telle que la RM ?

Nous avons vu dans le chapitre précédent que plus de quinze ans après les travaux précurseurs de Sharps et coll. (1993), différents auteurs (Martens et coll., 2006 ; M. S. McGlone & Aronson, 2006 ; Moè & Pazzaglia, 2006 ; Wraga et coll., 2006, 2007) se sont inspirés de la théorie de Steele pour montrer que la différence de sexe en RM pouvait être modulée par l'activation du stéréotype. Cette activation survient au cours du test, typiquement par le biais d'une manipulation des instructions. Cependant, nous estimons qu'en l'état, ces résultats ne sont pas suffisants pour accepter définitivement l'idée que la différence de sexe en RM est expliqué uniquement par une intervention du stéréotype. Nous pouvons regrouper les éléments qui affaiblissent la démonstration de ces études selon trois dimensions.

Tout d'abord, méthodologiquement, ces études ne permettent pas toujours une comparaison directe des résultats avec ceux de la littérature (p. ex. utilisation d'une version différente du test de RM, Wraga et coll., 2006, 2007). Mais surtout, les plans expérimentaux choisis ne permettent pas toujours non plus la comparaison entre les sexes au sein même de l'étude. En effet, souvent seules des femmes ont été testées (Wraga et coll., 2007) ou alors les hommes et les femmes n'étaient pas testés simultanément dans la même étude (Moè & Pazzaglia, 2006).

Plus important, lorsque la comparaison entre les sexes est possible, soit l'analyse statistique des résultats montre que la neutralisation du stéréotype conduit à une baisse des performances des hommes jusqu'au niveau des femmes (Wraga et coll., 2006), soit seule une variation de la performance des femmes est mise à jour, sans impact statistique sur la différence entre les deux sexes (Hausmann et coll., 2009 ; M. S. McGlone & Aronson, 2006). Or si la cause de la contre-performance des femmes est lié à l'activation d'un stéréotype, la neutralisation de ce stéréotype ou du moins sa « non-activation » devrait conduire les femmes à exprimer pleinement leur capacité, et atteindre un niveau comparable à ce-

lui des hommes. Pourtant l'on observe plutôt dans ces études un mouvement inverse avec une modulation de la performance des hommes, ceux-ci réalisant une contre-performance dans la condition *a priori* plus favorable aux femmes, ou à la rigueur non-menaçante. Ils n'exprimeraient alors pas leur plein potentiel, qu'ils atteindraient dans les conditions défavorables aux femmes. De tels résultats pourraient expliquer la réduction de la différence de sexe dans la condition contre-stéréotypique par un défaut de motivation de la part des hommes. En effet ils sont placés dans une situation de concurrence avec des femmes, mais la neutralisation du stéréotype ne leur permet pas de percevoir la comparaison avec ces femmes d'une manière qui leur serait favorable. En tout cas si régulation sociale il y a *via* l'activation (ou non) du stéréotype, il ne s'agit pas d'une régulation de la performance *des femmes* spécifiquement, comme le prévoit à première vue la théorie.

Enfin, la principale critique qui peut être formulée à l'encontre des travaux de Martens et coll. (2006), de Moè et Pazzaglia (2006) et de Wraga et coll. (2006, 2007) — et qui peut éventuellement expliquer le problème précédent — est la manière dont l'information sur le stéréotype est délivrée au sujet (c.-à-d. le type d'induction). Dans ces travaux les auteurs ne se contentent pas d'évoquer ou de falsifier l'idée d'une différence de sexe, ils précisent le sens de cette différence : les sujets apprennent que les femmes sont soit inférieures soit supérieures aux hommes sur le test proposé, et délivrent donc du même coup un feedback positif ou négatif sur le groupe d'appartenance. Cette manipulation n'est pas conforme à l'esprit du paradigme de départ (cf. Spencer et coll., 1999), toute variation des performances dans ces conditions est susceptible de refléter davantage l'effet d'un feedback (quelle qu'en soit la nature) positif ou négatif plus que l'intervention du stéréotype dont on fait l'hypothèse en tant que tel. Les conditions créées par ces auteurs imposent en réalité aux sujets un pronostic qui pourrait soit les encourager soit au contraire les conduire à se désengager de la tâche, or il s'agit plutôt de créer des conditions minimales propices à l'activation implicite ou explicite du stéréotype ou à sa neutralisation.

3.1 Objectifs de l'étude 1

Seules les études de M. S. McGlone et Aronson (2006), d'Hausmann et coll. (2009) et dans une certaine mesure de Sharps et coll. (1994) proposent des inductions correctes au regard de la théorie, mais ces études n'ont pu établir statistiquement que les femmes pouvaient atteindre les performances des hommes en RM.

Deux autres éléments concernant la population choisie et les conditions de passation du test participent à l'originalité de notre étude. Premièrement, il s'agit de montrer que l'effet du stéréotype peut intervenir dès l'adolescence, à un moment où le sexe devient un facteur d'identité particulièrement prégnant, d'autant que c'est précisément à cet âge que la différence de sexe à l'avantage des garçons en RM devient clairement mesurable. Parmi les rares études qui se sont intéressées à l'interférence du stéréotype chez les enfants et les adolescents, Ambady, Shih, Kim, et Pittinsky (2001) sont les premiers à notre connaissance à avoir montré la perméabilité de jeunes collégiennes à un stéréotype de sexe. De jeunes adolescentes asiatiques montraient une baisse de leur performance en mathématique lorsque leur identité de fille était activée (stéréotype négatif), mais à l'inverse de meilleures performances lorsque leur identité asiatique (stéréotype positif) était activée. Muzzatti et Agnoli (2007) en mobilisant des effectifs importants ont montré que les jeunes filles étaient sensibles à l'activation du stéréotype qui modifiait alors leur attitude vis-à-vis des mathématiques, et cela dès l'école élémentaire ($N = 476$) et au collège ($N = 271$).

Plus récemment Huguet et Régner (2007, 2009) ont reproduit des résultats qui montrent que le stéréotype négatif visant les femmes en mathématiques pouvait affecter la performance de collégiennes à un test visuo-spatial (figure complexe de Rey modifiée). Ces auteurs ont activé le stéréotype assez finement, par une simple présentation du test comme diagnostique des aptitudes en mathématiques (géométrie). Encourageants, ces résultats qui montrent l'importance de la perception de la tâche et des croyances afférentes pourraient expliquer la versatilité des différences obtenus avec le test de la figure complexe de Rey. Les données de Huguet et Régner (2007, 2009) indiquent que le traitement visuo-spatial chez les enfants n'est pas imperméable à l'interférence d'un stéréotype, et ce, dans

des conditions écologique proche du quotidien scolaire (classes entières). Ce dernier aspect, écologique, constitue un autre élément original de notre étude. Les questions de validité externe, souvent envisagées en terme de « conséquences pratiques » du phénomène, sont capitales dans ce domaine, tant la demande sociale est forte, notamment en termes d'intervention pour contrer le phénomène, ne serait-ce que dans le milieu scolaire.

Notre objectif étant d'établir distinctement le rôle du stéréotype sur les performances comparées des deux sexes en RM, il a alors été nécessaire de construire très précisément la situation de test en parant aux limitations soulevées précédemment. D'une part nous avons adopté un plan expérimental permettant des comparaisons multiples entre les sexes dans les différentes conditions (activation du stéréotype, falsification et condition contrôle, cette dernière étant souvent négligée), et d'autre part nous avons choisi des inductions respectueuses du paradigme original de falsification (Spencer et coll., 1999), laissant au sujet la latitude de « décider » du sens de la différence de sexe lorsque cette différence est évoquée.

Mais surtout, notre étude se distingue essentiellement par la taille de l'échantillon mobilisé, très importante ($N = 877$), afin de pouvoir montrer que la sensibilité au *stereotype threat* des filles dépend de leur niveau de performance en RM. Comme nous l'avons déjà souligné, il est aujourd'hui largement reconnu dans la littérature sur les différences de sexe que les analyses portant sur les tendances centrales — généralement la comparaison de la moyenne des distributions de performance des deux sexes — sont insuffisantes pour décrire correctement ces différences. D'une part, il est bien connu et accepté que les femmes et les hommes ne montrent pas la même variabilité dans leur performance (Feingold, 1992). Mais surtout, il est frappant de constater que les différences de sexe ne sont pas homogènes le long des distributions de performance : elles s'accroissent aux extrémités de ces distributions, ainsi le nombre d'hommes qui atteignent les plus hauts scores dépasse largement le nombre de femmes à ce niveau de performance (Hedges & Nowell, 1995 ; Halpern et coll., 2007).

Or la théorie du *stereotype threat* constitue un apport capital sur cette question, car elle permet très bien d'expliquer cette sur-représentation masculine dans les hauts niveaux de performance. En effet, selon cette théorie, les sujets les plus sensibles à l'interférence du stéréotype sont les individus ciblés les plus identifiés au domaine testé, qui se révèlent dans les faits être globalement les individus ciblés les plus performants dans le domaine. Pourtant, aucune étude sur le *stereotype threat* n'a cherché à réunir des données détaillées quant à la relation entre niveau de performance et impact de l'interférence. Tout au plus quelques études ont montré que le *stereotype threat* survenait typiquement chez des sujets d'un bon niveau dans le domaine testé (p. ex. des étudiants noirs d'un bon niveau universitaire concernant les tests d'intelligence, cf. Steele & Aronson, 1995, ou des étudiantes possédant de bonnes aptitudes mathématiques, cf. Spencer et coll., 1999). Mais aucun auteur n'a étudié systématiquement, comme nous le proposons, le rôle du niveau de performance dans l'expression de l'interférence du stéréotype. Cela peut s'expliquer par la lourdeur d'une telle étude : calculer une distribution de performance nécessite des échantillons de très grande taille, or les études dans ce domaine — dont les plans expérimentaux sont en général déjà assez contraignants — n'ont jamais porté que sur des échantillons de taille trop réduite pour ce genre d'analyses. Aussi l'étude que nous proposons dans ce chapitre se distingue nettement par son ampleur, dans le but d'offrir une finesse d'analyse inédite du phénomène d'interférence du stéréotype.

3.2 Méthode

3.2.1 Participants

L'un des objectifs de cette première étude étant de déterminer l'évolution de la proportion hommes / femmes (*sexe-ratio*) présents à différents niveaux de performance en fonction des conditions d'activation du stéréotype, un grand nombre de participants était

donc nécessaire. Ainsi, le protocole expérimental a été proposé à 877 collégiens (446 filles ; 431 garçons) de différents collèges du sud-est de la France (Bouches-du-Rhône, Hautes-Alpes et Drôme) de niveau 4^e ou 3^e, ce qui correspond à un âge moyen de $M = 14.04$ ($S.D. = .93$). Selon les données des académies concernées (Aix-Marseille et Grenoble), aucun des établissements choisis n'était classé ZEP, ni dans les catégories « violence » ou « sensible ». Une autorisation parentale de participation à l'expérience a été recueillie par le biais des établissements scolaires concernés.

3.2.2 Matériel

Le Test de Rotation Mentale

Le MRT de Vandenberg et Kuse a déjà été présenté de manière relativement détaillée dans notre premier chapitre (cf. p. 26). Ce test « papier-crayon » est divisé en deux sections (cf. annexe, première section p. 225 et seconde section p. 227) d'égale difficulté, chacune composée de dix problèmes soit vingt problèmes en tout. Pour chaque problème, le sujet doit sélectionner parmi quatre propositions deux représentations correctes d'un objet vu sous différents angles, dont un modèle est préalablement présenté (cf. fig 1.3, p. 26).

La méthode de cotation a été établie de manière à éliminer les réponses au hasard. Si une seule représentation incorrecte est sélectionnée par le sujet, alors l'ensemble du problème est compté faux, ce dont les participants, dans notre cas, ont été informés au début de la tâche. Si les deux représentations correctes sont sélectionnées, on compte deux points pour le problème ainsi résolu, et un point si le sujet n'a sélectionné qu'une seule réponse correcte (et seulement celle-ci). Tous les autres cas de figure ne rapportent aucun point. Le score maximal possible est donc de deux points par problème soit 40 points au total. La taille de l'effet de sexe selon cette méthode de cotation est typiquement de $d = .70$ (Voyer et coll., 1995).

Pour montrer une différence de sexe, le MRT doit être résolu en temps limité, classiquement cinq minutes par section, soit dix minutes au total. Vandenberg et Kuse proposent de baisser cette durée à 2×3 minutes pour les sujets jeunes (lycéens et étudiants), dont les performances sont supérieures à celles des autres groupes d'âges (les meilleures performances sont observées autour de 25 ans). Bien que la population testée était plus jeune (collégiens), nous avons adopté cette limite temporelle de trois minutes par section. Cela nous permettait d'une part d'exacerber la différence de sexe, et d'autre part de comparer les résultats des collégiens avec ceux généralement présentés dans la littérature concernant des sujets jeunes. À titre indicatif, on peut calculer à partir des résultats publiés par Albaret et Aubert (1996), concernant des lycéens français âgés de 15 ans, une taille d'effet de sexe de $d = .67$ lorsque la limite de temps est de trois minutes¹.

Questionnaire post-test

Ce questionnaire (cf. annexe p. 229), administré après le MRT afin de ne pas biaiser la manipulation de l'activation du stéréotype avant la mesure des performances en RM, était constitué de trois grands ensembles d'item.

Le premier ensemble de questions (cf. annexe p. 230) devait nous permettre d'établir le degré de connaissance du stéréotype d'infériorité des femmes par rapport aux hommes dans le domaine visuo-spatial. Une des dimensions forte des stéréotypes étant leur statut de croyance socialement partagée, connaître un stéréotype donné équivaut à reconnaître ce caractère partagé dans la société. Ainsi, pour chaque question, le sujet devait indiquer sur une échelle de Lickert combien de personnes, selon lui, partageaient une opinion (croyance) donnée en exemple. L'échelle allait de *1-personne* à *5-tout le monde*, et les croyances consistaient en l'idée d'une performance supérieure des hommes dans une si-

1. La taille d'effet de sexe atteint $d = .95$ si l'on prend en compte la totalité de l'échantillon de lycéens testé par ces auteurs, soit une population âgée de 15 à 19 ans, $N = 512$.

tuation donnée. Par exemple (première question) : *Y a-t-il des gens qui pensent que les femmes ont plus de mal que les hommes pour se servir d'une carte routière ?*

Le second ensemble de questions (cf. annexe p. 232) devait ensuite nous permettre d'établir le degré d'adhésion des participants au stéréotype. Il s'agissait donc du même ensemble de question que précédemment, mais le sujet devait indiquer sur une échelle allant de *1-pas du tout d'accord* à *5-tout à fait d'accord* son degré d'adhésion à l'idée proposée. Par exemple la première question devenait : *Personnellement, je pense que les femmes ont plus de mal que les hommes pour se servir d'une carte routière.*

Enfin le dernier ensemble (cf. annexe p. 234) était constitué de questions se rapportant à la perception du test par le sujet : difficulté perçue, auto-évaluation, auto-évaluation comparative, évaluation comparative des filles et des garçons (*manipulation check*), et évaluation comparative des sujets jeunes et âgés. Les participants devaient à chaque fois rapporter leur réponse sur une échelle de Lickert à cinq degrés.

3.2.3 Procédure

Les élèves ont été testés en condition écologique, c'est-à-dire en classe entière (mixte) dans leur salle de classe habituelle et en présence de leur professeur. La passation du test s'insérait dans l'emploi du temps des élèves en lieu et place d'un cours dont l'objet n'était ni les mathématiques, ni les sciences, ou encore le dessin. Ceci afin d'éviter que l'éventuelle appréhension de certains élèves vis à vis de ces matières n'interfère avec la perception spontanée du test, celle-ci étant un facteur critique de notre manipulation expérimentale. De même si un enseignement avait eu lieu l'heure précédant la passation, il ne concernait pas non plus ces disciplines. Une autre précaution concernant l'emploi du temps a été de choisir des horaires au sortir desquelles la communication entre les classes était peu probable, cela afin d'éviter tout échange concernant le test entre les sujets de classes différentes.

Deux catégories d'observables étaient recueillies : le score au MRT de Vandenberg et Kuse (1978), permettant d'évaluer la performance en rotation mentale 3D d'une part, et d'autre part les réponses au questionnaire conçu pour évaluer le positionnement et les croyances des élèves en rapport avec le stéréotype et la nature du test. Le questionnaire n'était distribué et rempli qu'après la passation du MRT afin que l'évocation du stéréotype dans les questions posées ne perturbe pas la manipulation expérimentale qui portait précisément sur l'activation de ce stéréotype.

L'expérimentateur, de sexe féminin ou masculin, expliquait dans un premier temps la manière dont le test allait se dérouler, et rappelait aux élèves que ce test n'était pas obligatoire et qu'ils pouvaient décider à tout moment de ne pas le poursuivre. Les différentes parties du test étaient ensuite distribuées aux élèves, qui devaient indiquer sur la première page leur date de naissance et leur sexe. La dernière phrase des instructions du MRT correspondait à l'induction de la condition expérimentale :

1. **standard** : instructions telles qu'elles sont habituellement délivrées aux élèves, sans aucune évocation explicite des performances, notamment en relation avec le sexe (cf. annexe p. 223) ;
2. **activation** : instructions standards auxquelles est ajoutée une dernière phrase, « *Pour information, en général, sur ce test, les garçons et les filles n'obtiennent pas les mêmes résultats ; ils n'ont pas la même performance.* ». Ici nous soulignons à nouveau que dans cette condition, le sens de la comparaison n'est pas explicité, c'est *a priori* la connaissance du stéréotype par les sujets qui suggère la supériorité des garçons à ce test spatial.
3. **falsification** : instructions standards auxquelles est ajoutée une dernière phrase qui falsifie explicitement le stéréotype : « *Pour information, en général, sur ce test, les garçons et les filles obtiennent les mêmes résultats ; ils ont la même performance.* »

Le plan expérimental était du type 2(Sexe) \times 3(Instructions) \times 2(Section du test). Les effectifs des élèves sont répartis comme suit entre les différentes conditions : 286 participants (150 filles, 136 garçons) dans la condition standard (contrôle), 283 (131 filles,

152 garçons) dans la condition d'activation explicite du stéréotype et 308 participants (165 filles, 143 garçons) dans la condition où les instructions falsifiaient le stéréotype.

Afin d'éviter une sur-réaction des élèves observée lors des pré-tests (commentaires à haute voix, polémiques, etc.), ces inductions n'étaient pas données à l'oral. Celles-ci n'apparaissaient qu'à l'écrit, à la fin des instructions, afin de manipuler l'activation du stéréotype chez les élèves au moment précis où ceux-ci étaient sur le point de s'engager dans le test de RM proprement dit. Par ailleurs, afin de limiter la suspicion des élèves quant à l'objet de la manipulation expérimentale, tous les élèves d'une classe donnée étaient placés dans une seule condition (*Standard*, *Activation* ou *Falsification*).

Le MRT est constitué de deux parties, qui doivent être résolues en temps limité (trois minutes chacune). Afin de maîtriser en classe entière les moments où les sujets se lançaient dans le test et s'arrêtaient, les instructions, de même que la première et la seconde section du MRT, étaient distribuées en trois feuillets séparés, face contre table. Les élèves ne retournaient chaque feuillet (dans l'ordre) qu'aux signaux correspondants. Ainsi, à un premier signal donné par l'expérimentateur, les élèves commençaient à lire les instructions du MRT, au sein desquelles trois problèmes de RM étaient donnés à résoudre dans un but de familiarisation (les solutions étaient livrées immédiatement après).

Une fois assuré que tous les élèves avaient lu les instructions et ne déclaraient rencontrer aucune difficulté de compréhension, l'expérimentateur donnait un signal invitant les sujets à commencer à résoudre la première partie du MRT. Au bout de trois minutes un second signal leur indiquait de s'arrêter. Après une pause d'une minute, les sujets disposaient à nouveau d'une durée de trois minutes — indiquées par le même dispositif — pour résoudre la seconde partie du test.

À l'issue du MRT, le questionnaire post-test était soumis aux élèves, auquel ils pouvaient répondre à leur rythme, sans pression temporelle. Il leur était toutefois précisé

oralement qu'il n'y avait pas de bonne ou de mauvaise réponse et qu'il leur fallait répondre de manière spontanée, selon la première pensée qui leur venait à l'esprit.

3.2.4 Attentes

Nos attentes portent sur différents niveaux d'analyse. L'un des principaux résultats de la littérature sur la différence de sexe est que cette différence s'exprime essentiellement par la sous-représentation des effectifs féminins dans les meilleurs niveaux de performance. Or, selon nous, le *stereotype threat* explique très bien la sous-représentation des filles dans les hauts scores. Pour montrer cet impact très localisé du stéréotype, nous proposons de quantifier l'asymétrie observée entre les effectifs masculins et féminins à différents niveaux de performances dans les différentes conditions, à l'aide d'une analyse spécifique : le calcul des sexe-ratios. Nous nous attendons alors à observer la classique sur-représentation des garçons pour les scores les plus élevés dans les conditions où le stéréotype est activé implicitement ou explicitement, c'est-à-dire dans les conditions *standard* et d'*activation* respectivement. Par contre nous nous attendons à une réduction très nette de cette asymétrie dans la condition de *falsification*, soit lorsque le stéréotype négatif est en principe neutralisé. D'où deux niveaux d'analyse de l'effet de nos conditions : une analyse menée sur l'échantillon global, et une autre analyse sur l'échantillon réduit aux meilleurs niveaux de performance, où le calcul des sexe-ratios aura révélé la localisation des effets du stéréotype sur la performance des filles.

Ainsi, sur l'échantillon pris dans son ensemble, nous ne nous attendons pas à pouvoir déceler un effet du stéréotype sur la performance des filles. Par contre, à ce niveau d'analyse, nous nous attendons à ce que l'effet de *stereotype lift* soit perceptible chez les garçons. Comme la falsification annule l'effet positif du stéréotype pour les garçons, nous devrions constater une performance plus basse des garçons dans cette condition par rap-

port aux deux autres conditions supposées activer le stéréotype. Nous attendons donc, sur l'échantillon total, une interaction d'ordre un : $Sexe \times Instruction$.

Par contre, sur l'échantillon composé des sujets les plus performants, nous nous attendons à déceler l'effet de *stereotype threat*, révélé par l'effet positif de la falsification du stéréotype sur les performances des filles, comparativement aux autres conditions, censées au contraire activer le stéréotype. Cet effet serait alors susceptible d'expliquer la sous-représentation féminine dans cette sous-partie de l'échantillon. Par ailleurs, du point de vu temporel, l'effet — c'est-à-dire l'annulation éventuelle de la différence de sexe lorsque les instructions falsifient le stéréotype — devrait être surtout perceptible dans la seconde section du MRT, car l'interférence trouverait sa source dans l'accumulation de pensées négatives au fur et à mesure que les filles se confronteraient à la difficulté du test (Cadinu et coll., 2005 ; Schmader et coll., 2008). Nous prévoyons donc de procéder à des analyses (ANOVA avec mesures répétées) dans le but de mettre à jour une interaction d'ordre deux : $Sexe \times Instruction \times Section$.

En somme, nous pensons que le *stereotype threat* est un phénomène plus difficile à mettre en évidence que le *stereotype lift*, car la théorie prévoit certes que les performances des filles sont entravées par l'intervention du stéréotype interférent, mais que cet effet est surtout sensible dans les hauts niveaux de performance, expliquant un phénomène analogue au fameux « plafond de verre ». Or ce dernier aspect est souvent négligé dans la littérature sur la RM, qui peine à montrer que l'activation du stéréotype empêche les femmes d'atteindre des scores aussi élevés que ceux obtenus par les hommes qui font preuve des meilleures aptitudes en RM. Nous avons discuté auparavant le fait que de nombreuses études publiées sur le stéréotype et les différences de sexe en RM ont surtout montré une influence du stéréotype sur la performance des hommes. De notre point de vue, ces études ont essentiellement mis en évidence une augmentation de la performance masculine, du fait d'un contexte favorable engendré par l'activation du stéréotype (*stereotype lift*). Nous sommes les premiers à nous donner les moyens de localiser un effet positif de la falsification

sur le score des filles dans les hauts niveaux de performance, effet qui les conduit à égaler les scores des garçons (et non l'inverse), jusqu'à annuler la différence de sexe. Nous avons pu atteindre cette finesse d'analyse grâce au recrutement d'un échantillon suffisamment grand, qui nous permet d'observer les effets du stéréotype spécifiquement chez les filles qui réalisent les plus hauts scores. Les études publiées précédemment ont mobilisé des échantillons trop petits qui ne permettaient pas une telle précision dans l'analyse de la distribution de performance : la mesure d'un effet du stéréotype sur la performance des participantes testées était donc hors de leur portée.

3.3 Résultats

3.3.1 Test de Rotation Mentale

Sexe-ratio

Dans un premier temps, nous cherchons à rendre compte de la manière dont le *stereotype threat* altère la forme de la distribution de performance au MRT des filles et des garçons, au niveau des hauts scores. L'analyse doit pouvoir montrer d'une part que le nombre de filles qui réalisent les scores les plus élevés est proportionnellement plus important dans la condition de *falsification* que dans les deux autres conditions — où la disparition classique des filles dans les hauts score doit également être observée. D'autre part, et dans des proportions moindres, les garçons doivent montrer le profil inverse (*stereotype lift*). Compte tenu de ces attentes, nous proposons un outil d'analyse qui permet de constater l'asymétrie de la répartition des effectifs féminins et masculins dans les hauts niveaux de performance en fonction des conditions. Il s'agit du calcul des sexe-ratios par tranche de performance, que nous décrivons ci-après.

Calcul des sexe-ratios. Nous proposons de diviser l'échantillon total en cinq sous-groupes (ou tranches) correspondants à cinq niveaux de performance. Pour une meilleure lisibilité nous qualifions ces niveaux de *très mauvais*, *mauvais*, *moyen*, *bon* et *très bon*. Le sous-groupe *très mauvais* est constitué par les 20% des participants des deux sexes dont les scores sont les plus faibles. La tranche *mauvais* par les 20% suivants, et ainsi de suite, jusqu'à la cinquième et dernière tranche (niveau *très bon*) composée des participant(e)s les plus performant(e)s. Pour chaque tranche, le calcul du sexe-ratio nous renseigne sur l'état de l'asymétrie entre les effectifs des deux sexes. Nous nous sommes inspiré de la procédure employée par Hedges et Nowell (1995), ces auteurs ayant calculé les sexe-ratios pour différents *percentiles* ($\geq 90\%$; $\geq 95\%$ et $\leq 10\%$). Ce découpage en percentile ne convient pas dans le cas présent. En effet, notre objectif est de comparer des rapports entre effectifs à différents niveaux de performance, alors que la comparaison des percentiles impose qu'une partie des individus soient inclus dans chaque percentile (par exemple les participants composant le percentile $\geq 95\%$ sont aussi inclus dans le percentile $\geq 90\%$) induisant un biais dans la comparaison. La formule correspondante aux calculs que nous présentons est :

$$sexe-ratio = \frac{\frac{\text{Nombre de garçons au niveau de performance considéré}}{\text{Nombre total de garçons}}}{\frac{\text{Nombre de filles au niveau de performance considéré}}{\text{Nombre total de filles}}}$$

Il découle de cette formule que, proportionnellement au nombre total de garçons et de filles, un sexe-ratio supérieur à 1 indique une sur-représentation des garçons, tandis qu'un sexe-ratio inférieur à 1 une sur-représentation des filles.

Résultats. L'analyse des distributions de performance (cf. fig. 3.1 p. 92) montre que, conformément à nos hypothèses, le nombre de garçons dépasse largement le nombre de filles dans la tranche des *très bons* scores dans les conditions *standard* (ratio = 3.42) et d'*activation* (ratio = 4.05). Par contre, lorsque le stéréotype est falsifié le nombre de

3.3 Résultats

garçons ne dépasse plus que modestement le nombre de filles (ratio = 1.59) (cf. fig. 3.1, p. 92).

En outre, cette augmentation de la proportion de filles dans la tranche des *très bons* suite à la *falsification* du stéréotype correspond bien à une augmentation du nombre de filles par rapport aux deux autres conditions (*standard et activation*), et non à une baisse du nombre de garçons. En effet nous pouvons voir dans la figure 3.1 que le nombre de filles dans la tranche des *très bons* double par rapport aux autres conditions tandis que le nombre de garçons est légèrement inférieur — la falsification annulant également le *stereotype lift* favorable aux garçons — mais dans de bien moindres proportions (environ -15%).

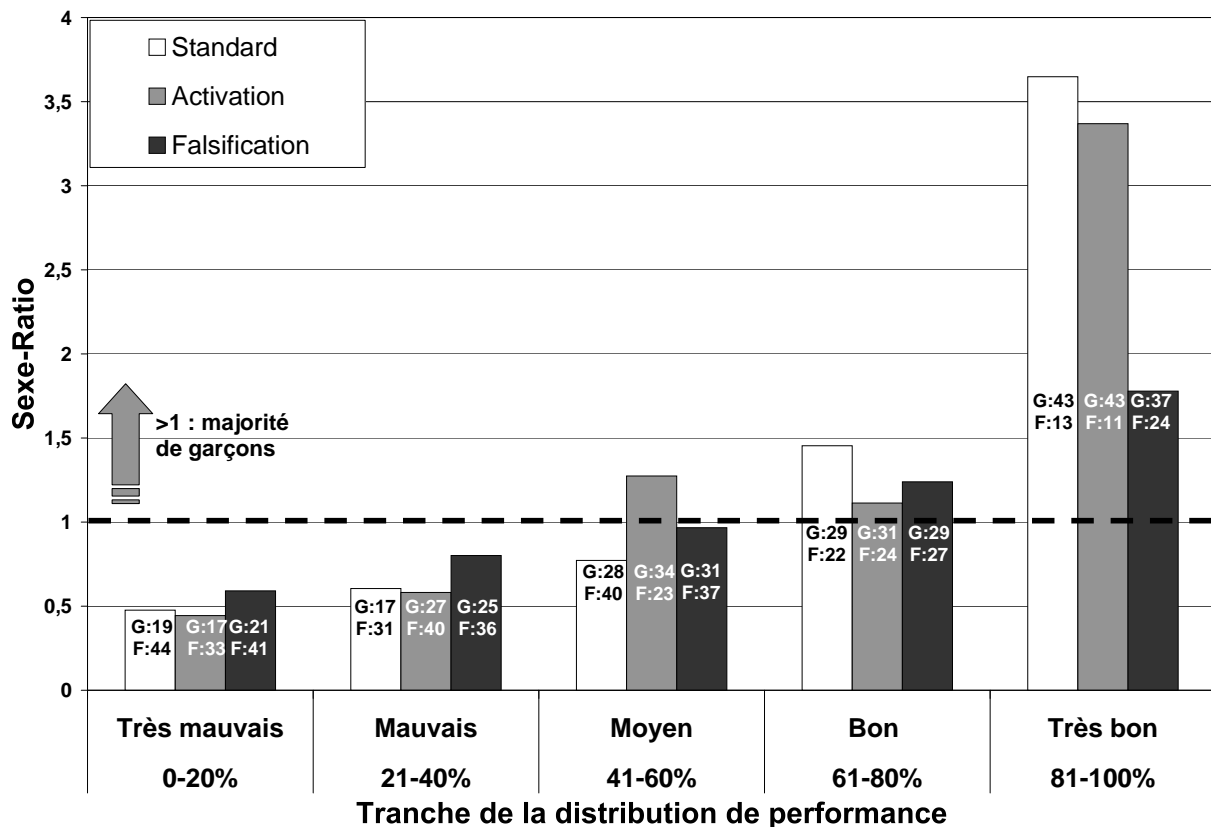


FIGURE 3.1 — Sexe-ratios pour différentes tranches de performance. Chaque tranche représente 20% de l'effectif total.

Ce résultat confirme que ce sont les filles obtenant les meilleures performances au MRT qui constituent la population féminine la plus sensible à l'interférence du stéréotype. On se

souvent que cette interférence est attendue plus forte dans la seconde section du test, c'est exactement ce que montrent les sexe-ratios calculés pour chacune des deux parties du test. Nous observons notamment en condition d'*activation* une diminution assez spectaculaire dans les plus haut scores du nombre de filles dans la seconde partie du test (cf. annexe p. 235, fig. A.1).

Nous procéderons donc dans la section suivante à une analyse classique des scores obtenus au MRT dans l'échantillon total, puis nous focaliserons nos analyses sur les effectifs pour lesquels la présente analyse des sexe-ratios a montré que les différences de sexes deviennent critiques. Ce sous-échantillon correspond au tiers supérieur de la distribution de performance ($N = 304$), ou en pratique à un « *cut-off* » à partir d'un score supérieur à 16.

Résultats sur la totalité de l'échantillon

Les scores au MRT ont été soumis à une analyse de variance mixte avec le sexe des participants et le type d'instruction en facteurs inter-sujets et la section du test (partie 1 vs. 2) en mesure répétée.

Effet des instructions et décours de la performance. Nous observons d'abord un effet classique de sexe, $F(1,871) = 80.1$, $p < .001$; $\eta_p^2 = .084$, indiquant que les filles ($M = 6.23$, $SD = 3.59$) produisent une performances inférieure à celle des garçons ($M = 8.43$, $SD = 3.74$). Un effet principal des instructions apparaît également : $F(2,871) = 5.12$, $p < .01$; $\eta_p^2 = .012$, les performances en *falsification* se révèlent globalement inférieures ($M = 6.78$, $SD = 3.69$) à celles observées dans les conditions *standard* ($M = 7.61$, $SD = 3.72$) et *activation* ($M = 7.60$, $SD = 3.70$). Mais surtout, nous observons une interaction *Sexe* \times *Instructions*, $F(2,871) = 3.31$, $p < .05$; $\eta_p^2 = .008$ (cf. fig. 3.2 p. 94), conforme à nos attentes et aux observations d'autres auteurs dans ce domaine

(cf. chapitre 2 p. 52). Pour approfondir cette interaction nous avons réalisé une analyse par contrastes orthogonaux opposant pour chaque sexe la condition de *falsification* à la moyenne des deux autres conditions. Cette analyse confirme que notre interaction est due à la performance supérieure des garçons dans les conditions *standard* ($M = 17.82$, $SD = 8.23$) et d'*activation* ($M = 17.87$, $SD = 8.49$), relativement à la condition de *falsification* ($M = 14.87$, $SD = 7.47$). Comme attendu, le score des filles s'avère identique dans toutes les conditions. En bref, nous observons sur la population globale un phénomène de *stereotype lift* : les garçons produisent une performance supérieure dans les deux conditions où le stéréotype est supposé activé (soit implicitement, soit explicitement) relativement à la condition où l'idée d'une différence de sexe est falsifiée d'emblée par les instructions. Cet effet s'exprime de la même façon dans les deux sections du test, aucune interaction en effet n'est observée avec ce dernier facteur, $F(2,871) < 1$. Cependant, nous constatons une baisse globale des scores entre la première ($M = 7.75$, $SD = 4.42$) et la seconde section ($M = 6.91$, $SD = 4.16$) du test, $F(1,871) = 34.00$, $p < .001$; $\eta_p^2 = .038$.

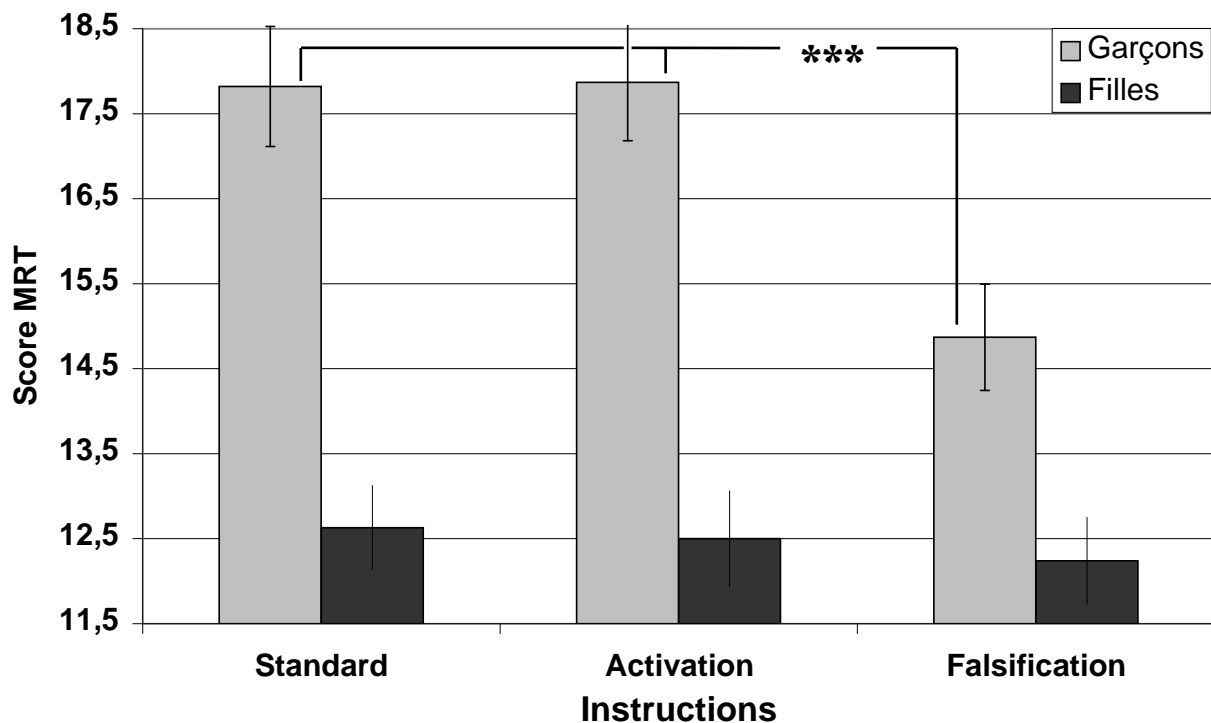


FIGURE 3.2 – Interaction *Sexe* \times *Instructions* sur l'échantillon total ($N=877$).

Taille des effets. Dans le domaine des différences de sexe, il est d’usage de quantifier la taille des effets de sexe à l’aide du d de Cohen². Ce calcul (cf. tableau 3.1 p. 95) montre que la différence entre les deux groupes de sexe dans les conditions *standard* et d’*activation* est de grande taille et identique aux valeurs rapportées par la plupart des études sur la RM. Cette différence est réduite d’un facteur deux dans la condition de *falsification*. Ceci étaye donc non seulement l’idée que l’intervention d’un stéréotype de sexe module l’expression de la différence en RM, mais également que des instructions sans référence explicite au stéréotype ne préviennent pas des effets de ce dernier. Les instructions standards sont neutres *a priori* et c’est ainsi qu’elles sont appréhendées dans la grande majorité des recherches sur la RM. Or, nos résultats montrent une similitude certaine entre cette condition soit-disant neutre et la condition d’activation explicite du stéréotype. Aussi, des instructions neutres à l’égard du stéréotype, si elles ne favorisent pas son intervention, ne l’empêchent pas non plus. Aucune trace du *stereotype threat* dans l’échantillon global, mais encore une fois ceci était attendu puisque précisément cet autre phénomène ne s’exprime en principe que chez les individus à la fois ciblés par le stéréotype mais aussi très performants dans le domaine considéré. Dans la section suivante, nous présentons les analyses concernant le sous-échantillon isolé par l’analyse des sexe-ratios, où nous nous attendions à un effet critique du stéréotype sur la performance des filles.

TABLE 3.1 – Taille de effet l’effet de sexe (d de Cohen) pour chaque section du MRT et pour l’ensemble du test, par type d’instruction (échantillon total, $N = 877$).

	Instructions		
	Standard	Activation	Falsification
$d_{Section\ 1}$.58	.64	.38
$d_{Section\ 2}$.69	.58	.27
d_{Total}	.72	.70	.37

2. J. Cohen (1992) propose de mesurer la taille d’une différence entre deux groupes A et B à l’aide de la formule : $d = \frac{M_A - M_B}{\sigma_{total}}$ qui est donc une mesure proportionnelle à l’écart-type. Cohen propose de qualifier d de *grand* lorsque $d \approx .80$, de *modéré* lorsque $d \approx .50$ et enfin de *petit* lorsque $d \approx .20$.

Collégien(ne)s montrant les meilleures capacités en RM

Les scores au MRT ont été soumis à une analyse de variance mixte avec le sexe des participants et le type d'instruction en facteurs inter-sujets et la section du test (partie 1 vs. 2) en mesure répétée.

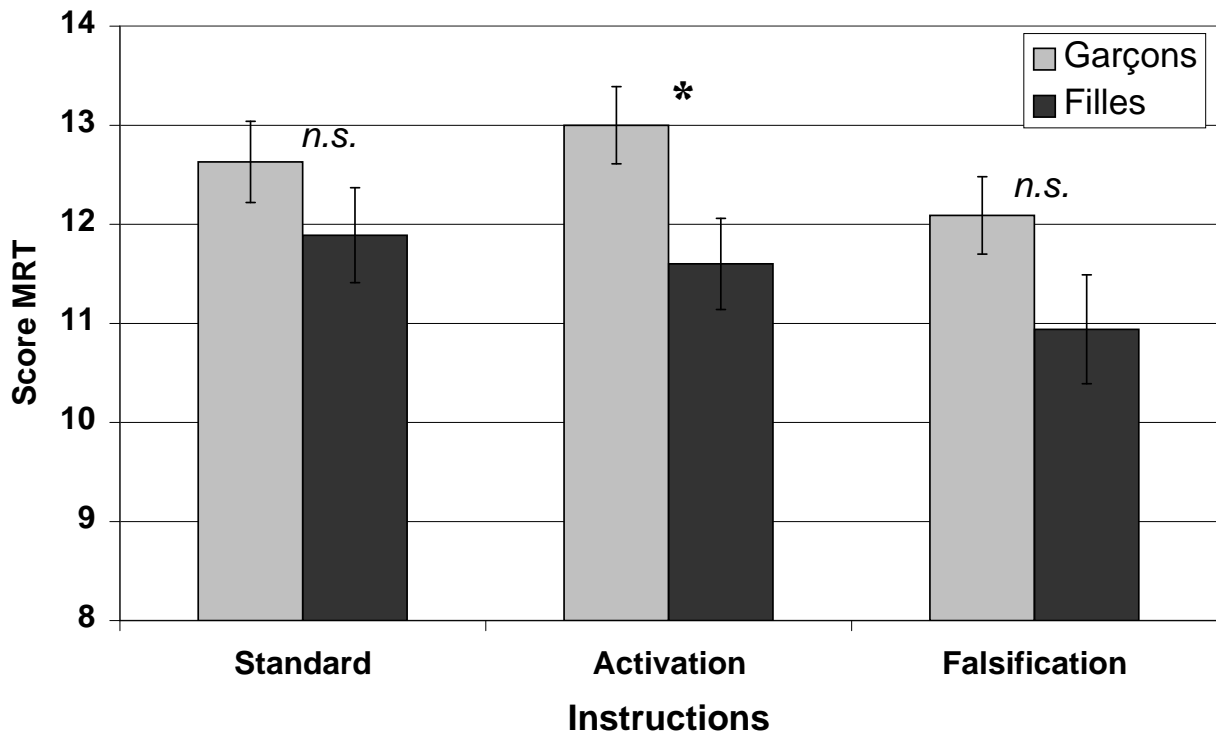
Décours de la performance. Nous prévoyions que l'effet des instructions sur la performance des collégiennes intervienne entre la première et la seconde section du MRT, car l'interférence du stéréotype chez les jeunes filles est plus importante au fur et à mesure qu'elles se confrontent à la difficulté du test. Cette hypothèse est confirmée par l'ANOVA mixte (cf. fig. 3.3-a et b p. 98)³ et conformément à notre hypothèse, l'interaction d'ordre deux $Sexe \times Instructions \times Section$, $F(2,298) = 4.24$, $p < .05$; $\eta_p^2 = .029$ est significative.

Dans la première section du test (fig. 3.3-a), les contrastes orthogonaux indiquent une différence significative entre les sexes seulement pour les instructions activant explicitement le stéréotype. Cela confirme donc notre hypothèse que l'importante différence de sexe observée sur ce test n'intervient pas dans la première partie du MRT chez les participants les plus doués, à moins de suggérer explicitement l'existence d'une telle différence dans les instructions. Par contre, dans la seconde section du test, les contrastes orthogonaux montrent qu'un écart apparaît entre les sexes dans toutes les conditions, sauf lorsque les instructions falsifient le stéréotype. Alors que la falsification conduit même de manière relativement inattendue les filles à réaliser une performance ($M = 11.27$, $SD = 6.81$) supérieure à celle des garçons ($M = 10.47$, $SD = 5.38$) — écart non significatif — cf. fig. 3.3-b, les contrastes révèlent une différence de sexe significative dans les conditions *standard* (filles : $M = 8.80$, $SD = 5.63$; garçons : $M = 11.35$, $SD = 5.71$) et d'*activation* (filles : $M = 9.49$, $SD = 5.72$; garçons : $M = 12.04$, $SD = 5.18$).

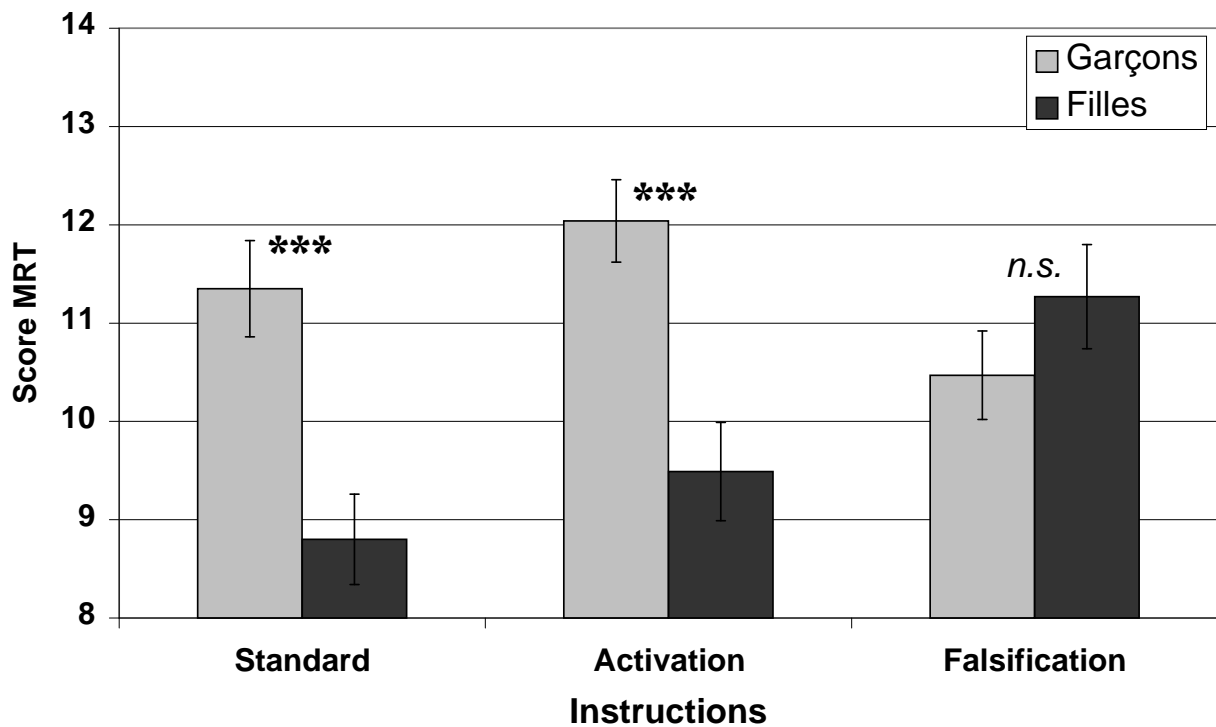
3. Afin de permettre la comparaison avec l'échantillon total, un graphique similaire montrant les scores obtenus aux deux sections du test pour l'échantillon global est disponible en annexe, fig. A.2 p. 236

Ainsi, conformément à nos attentes, ces résultats indiquent que la différence de performance due au stéréotype apparaît dans la seconde section du test, une fois que les participantes se sont suffisamment confrontées à la difficulté du test. Par comparaison avec les résultats observés sur l'échantillon total (cf. annexe p. 236 fig. A.2-a et b), nous constatons que le décours des performances est spécifique aux niveaux élevés de performance.

Pour mieux préciser l'effet des instructions dans la seconde section du test, nous avons également réalisé des contrastes opposant la condition de *falsification* à la moyenne des deux autres conditions pour chaque sexe. Il apparaît que la performance des garçons est supérieure dans les conditions où le stéréotype est supposé activé par rapport à la condition de *falsification*, ce qui indique une effet de *stereotype lift*. La performance des filles est quant à elle inférieure dans ces deux conditions à celle observée dans la condition de *falsification*, ce qui indique dans leur cas un effet de *stereotype threat*. Ces comparaisons ne sont pas significatives dans la première section du test, ce qui confirme encore notre hypothèse sur la dynamique de l'effet, qui n'apparaît qu'une fois que les participants se sont suffisamment confrontés à la difficulté du test.



(a) Première partie du MRT



(b) Seconde partie du MRT

*** : $p < .001$; ** : $p < .01$; * : $p < .05$

FIGURE 3.3 – Interaction *Sexe* \times *Instructions* \times *Section* (score > 16).

Taille des effets. Si l'on considère les scores à l'ensemble du test, l'écart entre les sexes est *grand* dans les conditions *standard* ($d = .64$) et d'*activation* ($d = .80$), une taille d'effet du même ordre que celle rapportée classiquement pour ce test (Voyer et coll., 1995). Par contre la taille de l'effet est proche de zéro dans la condition de *falsification* ($d = .07$), indiquant une disparition de l'effet de sexe (cf. tableau 3.2 p. 100)!

Pour prendre la mesure du décours des performances, nous avons calculé la taille des effets de sexe dans chaque condition pour les deux sections du test. Dans la première section, les différences de sexe sont *modérées* dans les conditions d'*activation* ($d = .44$) et de *falsification* ($d = .38$). Enfin, l'effet de sexe est *petit* ($d = .22$) lorsque les sujets reçoivent des instructions standards. Ces calculs corroborent nos précédents résultats : ce n'est pas dans la première partie du test que la différence est la plus importante, même si dans la condition de falsification la taille modérée de la différence (en faveur des garçons) peut poser question. L'effet de sexe n'est pas significatif pour autant dans cette condition, tout au plus peut-on imaginer que la falsification n'a pas poussé les filles à fournir un effort conséquent dès le début du test. Dans la seconde section, conformément à nos attentes, la divergence entre les tailles des effets apparaît très nettement. Ainsi, $d = .69$ dans la condition *standard* et $d = .75$ dans la condition d'*activation*, ce qui correspond à l'importante taille d'effet classiquement observée pour ce test. À l'inverse, la *falsification* du stéréotype provoque une inversion du sens de la différence de sexe ($d = -.25$). Bien que non-significative, cette inversion est suffisamment rare — à notre connaissance aucune manipulation expérimentale n'a jamais pu en provoquer — pour être signalée. L'analyse de la taille des effets de sexe constitue donc un indice supplémentaire en faveur de notre hypothèse sur la localisation de l'effet du stéréotype essentiellement dans la seconde section du test.

TABLE 3.2 – Taille de l'effet de sexe (d de Cohen) pour chaque section du MRT et pour l'ensemble du test, par type d'instruction (échantillon limité aux meilleurs scores, $N = 304$).

	Instructions		
	Standard	Activation	Falsification
$d_{Section\ 1}$.22	.44	.38
$d_{Section\ 2}$.69	.75	-.25
d_{Total}	.64	.80	.07

3.3.2 Questionnaire

Les réponses à chaque question ou ensemble de questions ont été soumises à une analyse de variance avec le sexe des participants et le type d'instruction en facteurs.

Croyance en l'existence du stéréotype⁴

Échantillon Total. On se souvient qu'avec cette échelle, les sujets devaient répondre à des questions du type *Y a-t-il des gens qui pensent que les femmes ont plus de mal que les hommes pour se servir d'une carte routière ? (1) Personne / ... / (5) Tout le monde*. Cette échelle était conçue pour savoir dans quelle mesure nos participants avaient connaissance de l'existence d'un stéréotype stipulant l'infériorité des femmes par rapport aux hommes dans le domaine visuo-spatial. Cette échelle est fiable avec un indice de Cronbach $\alpha = .78$. Globalement, les participants montrent une légère résistance à admettre une telle croyance⁵ ($M = 2.69$, $SD = .77$ sur une échelle de 1 à 5). L'ANOVA révèle un effet principal du sexe, $F(1,871) = 41.69$; $p < .001$; $\eta_p^2 = .046$. Les filles résistent plus à cette croyance ($M = 2.53$, $SD = .64$) que les garçons ($M = 2.86$, $SD = .84$). Les effets de la condition, $F(2,871) < 1$, et l'interaction *Sexe* \times *Instruction*, $F(2,871) < 1$, sont non significatifs.

4. Nous rappelons que l'ensemble des questions constituant l'échelle se trouve en annexe p. 230

5. Un test-t pour un échantillon indique que la moyenne de l'échelle de croyance est significativement inférieure à 3, $t(876) = -11.82$, $p < .001$.

Échantillon Réduit. Avec un $\alpha = .80$, l'échelle de connaissance du stéréotype est également fiable pour l'échantillon restreint aux élèves montrant les meilleurs scores au MRT. Les participants montrent également une légère résistance à admettre cette croyance⁶ ($M = 2.71$, $SD = .78$). L'ANOVA montre un effet principal de sexe, $F(1,298) = 8.83$; $p < .005$; $\eta_p^2 = .029$, ce qui une nouvelle fois indique une moindre résistance des garçons ($M = 2.79$, $SD = .83$) par rapport aux filles ($M = 2.52$, $SD = .65$) à l'égard de cette croyance. Les effets des instructions, $F(2,298) = 1.79$, et de l'interaction *Sexe* \times *Instruction*, $F(2,298) < 1$, ne sont pas significatifs.

En conclusion, les filles, quel que soit l'échantillon considéré (global vs. réduit), admettent plus difficilement l'existence du stéréotype, par rapport aux garçons.

Adhésion au stéréotype⁷

Échantillon Total. On se souvient qu'avec cette échelle, les sujets devaient répondre à des questions du type *Personnellement, je pense que les femmes ont plus de mal que les hommes pour se servir d'une carte routière. (1) Pas du tout d'accord / ... / (5) Tout à fait d'accord*. Cette échelle était conçue pour savoir dans quelle mesure nos participants adhéraient eux-mêmes au stéréotype stipulant l'infériorité des femmes par rapport aux hommes dans le domaine visuo-spatial. L'échelle d'adhésion au stéréotype est fiable, $\alpha = .78$. La corrélation est élevée avec l'échelle précédente : $\rho = .63$, $p < .001$. En fait on peut considérer que la première échelle mesure une dimension plus implicite de la croyance (inférence sur la diffusion de cette croyance dans la société), alors que l'échelle d'adhésion est plus explicite au sens où elle mesure une croyance propre au sujet et déclarée comme telle. Les participants rapportent dans l'ensemble une résistance plus forte encore (par rapport à l'échelle précédente) pour adhérer au stéréotype⁸ ($M = 2.50$, $SD = 1.01$).

6. Un test-t pour un échantillon indique que la moyenne de l'échelle de croyance est significativement inférieure à 3, $t(303) = -6.55$, $p < .001$.

7. Nous rappelons que l'ensemble des questions constituant l'échelle se trouve en annexe p. 232

8. Un test-t pour un échantillon indique que la moyenne de l'échelle d'adhésion est significativement inférieure à 3, $t(876) = -14.56$, $p < .001$.

L'ANOVA montre un effet principal de sexe, $F(1,871) = 127.54$, $p < .001$; $\eta_p^2 = .13$, les filles déclarent rejeter plus fortement le stéréotype ($M = 2.15$, $SD = .73$) que ne le font les garçons ($M = 2.87$, $SD = 1.12$). Les effets des instructions, $F(2,871) = 2.14$, et de l'interaction $Sexe \times Instructions$, $F(2,871) = 2.74$, ne sont pas significatifs.

Échantillon Réduit. L'échelle d'adhésion reste fiable pour l'échantillon sélectif, $\alpha = .72$, avec encore une fois le même profil de résultats : dans l'ensemble les participants déclarent peu adhérer au stéréotype⁹ ($M = 2.51$, $SD = 1.13$) et l'ANOVA montre seulement un effet principal du sexe, $F(1,298) = 35.97$, $p < .001$; $\eta_p^2 = .11$. Les filles ($M = 1.99$, $SD = .66$) rejettent plus fortement le stéréotype que les garçons ($M = 2.78$, $SD = 1.22$). Les effets des instructions, $F(2,298) = 1.89$, et de l'interaction $Sexe \times Instructions$, $F(2,298) = 1.48$, ne sont pas significatifs. Les filles déclarent donc moins adhérer que les garçons au stéréotype.

Perception du test

Les questions suivantes devaient nous permettre de déterminer comment le test était perçu (difficulté perçue, auto-évaluation, auto-évaluation comparative, etc.). Ce questionnaire était nécessairement soumis aux participants après le MRT, et leur réussite à ce test a pu influencer les réponses à ces questions. Aussi, nous indiquons non seulement les résultats de l'ANOVA avec le sexe et le type d'instructions en facteurs, mais aussi les résultats de l'analyse lorsque le score obtenu au MRT est introduit en covarié.

Nous verrons que les analyses sur les questionnaires conduisent aux mêmes conclusions pour l'échantillon total et l'échantillon réduit, mais pour des raisons de clarté et de précision, ces analyses sont présentées systématiquement pour ces deux échantillons pour chaque étape du questionnaire.

9. Un test-t pour un échantillon indique que la moyenne de l'échelle d'adhésion est significativement inférieure à 3, $t(303) = -7.50$, $p < .001$.

Estimation de la difficulté du test

Pour mémoire, les participants devaient donner une estimation de la difficulté perçue du test en répondant à la question : *Pour toi ce test était : (1) très difficile / ... / (5) très facile.*

Échantillon Total. Dans l'ensemble, le test est perçu comme plutôt difficile¹⁰ ($M = 3.24$, $SD = .98$ sur une échelle de 1 à 5). L'ANOVA révèle un effet de sexe, $F(1,871) = 25.10$, $p < .001$; $\eta^2 = .028$, les filles jugent le test plus difficile ($M = 3.08$, $SD = .93$) que les garçons ($M = 3.41$, $SD = 1.00$). L'analyse révèle par ailleurs un effet principal des instructions, $F(2,871) = 3.25$, $p < .01$; $\eta_p^2 = .007$, les élèves soumis aux instructions *standards* jugent le test plus difficile ($M = 3.15$, $SD = .95$) que ceux exposés aux instructions *d'activation* ($M = 3.37$, $SD = .98$) et de *falsification* ($M = 3.21$, $SD = 1.00$). L'interaction $Sexe \times Instructions$, $F(2,871) < 1$, est non significative. Les mêmes effets sont observés suite à l'introduction du score au MRT en covarié (effet des *Instructions*, $F(1,870) = 3.17$, $p < .05$; $\eta^2 = .007$ et effet du *Sexe*, $F(1,870) = 17.00$, $p < .001$; $\eta^2 = .019$).

Échantillon Réduit. Dans l'ensemble, le test est également perçu comme plutôt difficile¹¹ ($M = 3.36$; $SD = .94$) par les sujets qui obtiennent les meilleurs scores. Seul l'effet de sexe est significatif, $F(1,871) = 11.13$, $p < .001$; $\eta_p^2 = .036$, les filles jugent le test de difficulté assez modérée ($M = 3.12$, $SD = .88$), alors que les garçons ont tendance à le juger plus facile ($M = 3.49$, $SD = .95$). Les effets des instructions, $F(2,298) = 1.50$ et de l'interaction $Sexe \times Instructions$, $F(2,298) = 1.67$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 5.31$, $p < .05$; $\eta^2 = .018$. Cette dernière analyse met également à jour une interaction d'ordre un $Instructions \times Sexe$, $F(1,870) = 3.07$, $p < .05$; $\eta^2 = .020$), les filles jugeant le test

10. Un test-t pour un échantillon indique que la moyenne de la difficulté perçue est significativement supérieure à 3, $t(876) = 7.32$, $p < .001$.

11. Un test-t pour un échantillon indique que la moyenne de la difficulté perçue est significativement supérieure à 3, $t(303) = 6.65$, $p < .001$.

particulièrement plus difficile dans les conditions *Standard* ($M = 3.17$, $SD = .15$) et *Falsification* ($M = 2.95$, $SD = .16$). Les filles, malgré leurs bons résultats, jugent donc le test plus difficile que les garçons au même niveau de performance.

Auto-évaluation en RM

Pour mémoire, l'auto-évaluation était mesurée en posant la question suivante : *Comment te juges-tu sur ce test : (1) très mauvais(e) / ... / (5) très bon(ne)*.

Échantillon Total. Dans l'ensemble, les participants s'auto-évaluent plutôt positivement¹² ($M = 3.10$, $SD = .93$) mais ce phénomène est lié essentiellement aux garçons comme le montre l'effet de sexe lui aussi significatif, $F(2,871) = 32.11$, $p < .001$; $\eta_p^2 = .036$: les filles s'auto-évaluent moins positivement ($M = 2.85$, $SD = .85$) que ne le font les garçons ($M = 3.28$, $SD = .97$). Les effets des instructions, $F(2,871) = 1.19$ et de l'interaction *Sexe* \times *Instructions*, $F(2,298) < 1$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 22.89$, $p < .001$; $\eta^2 = .026$. Non seulement à performances égales les filles se jugent moins performantes que les garçons, mais ce jugement contraste avec leurs déclarations explicites à l'égard du stéréotype.

Échantillon Réduit. Nous observons toujours une différence de sexe, $F(1,298) = 14.93$, $p < .001$; $\eta_p^2 = .048$, due au fait que les filles se jugent juste moyennes au test ($M = 2.99$, $SD = .69$) alors que les garçons s'estiment plutôt bons ($M = 3.37$, $SD = .86$). Les effets des instructions, $F(2,298) < 1$, et de l'interaction *Sexe* \times *Instructions*, $F(2,298) < 1$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 11.77$, $p < .001$; $\eta^2 = .038$. Nous observons donc la même

12. Un test-t pour un échantillon indique que la moyenne de l'auto-évaluation est significativement supérieure à 3, $t(876) = 3.32$, $p < .001$.

contradiction que sur l'échantillon total : les filles rejettent le stéréotype mais expriment néanmoins une certaine difficulté à s'auto-évaluer positivement.

Auto-évaluation Comparative

Pour mesurer comment les participants s'auto-évaluent comparativement aux autres élèves, la question suivante leur était posée : *Sur ce test comment te juges-tu par rapport aux autres élèves de ta classe : (1) beaucoup moins bon(ne) / ... / (5) bien meilleur(e).*

Échantillon Total. Les élèves estiment globalement leur propre performance égale à celle des autres élèves de leur classe¹³ ($M = 2.95$, $SD = .77$). L'ANOVA révèle un effet de sexe, $F(1,871) = 43.19$, $p < .001$; $\eta_p^2 = .047$, les filles déclarent que cette comparaison leur est plutôt défavorable ($M = 2.79$, $SD = .67$), alors que les garçons se jugent légèrement supérieurs aux autres élèves ($M = 3.13$, $SD = .83$). Les effets des instructions, $F(2,871) < 1$, et de l'interaction *Sexe* \times *Instruction*, $F(2,871) < 1$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 27.97$, $p < .001$; $\eta^2 = .031$. Donc, de manière cohérente, non seulement les filles jugent leur performance inférieure à la moyenne dans l'absolu, mais elles se jugent aussi légèrement inférieure aux autres élèves (la corrélation est significative entre l'auto-évaluation et l'auto-évaluation comparative, $\rho = .40$; $p < .001$), et ce, même si l'on contrôle leur performance obtenu au MRT. Ce jugement est toujours en contradiction avec leur résistance à admettre le stéréotype (au moins dans leurs déclarations).

Échantillon Réduit. Les sujets obtenant les meilleures performances estiment globalement leur propre performance très légèrement supérieure à celle des autres élèves

13. Un test-t pour un échantillon indique que la moyenne de l'auto-évaluation comparative n'est pas significativement différente de 3, $t(876) = -1.75$, *n.s.*

de leur classe¹⁴ ($M = 3.09$, $SD = .71$). L'ANOVA montre toujours un effet de sexe, $F(1,298) = 18.00$, $p = .001$; $\eta_p^2 = .057$: même les filles les plus performantes se perçoivent défavorablement par rapport aux autres élèves ($M = 2.84$, $SD = .59$) alors qu'à l'inverse les garçons jugent que cette comparaison leur est favorable ($M = 3.21$, $SD = .73$). Les effets des instructions, $F(2,871) < 1$, et de l'interaction $Sexe \times Instruction$, $F(2,871) = 1.08$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 13.08$, $p < .001$; $\eta^2 = .042$. Le fait de mieux réussir le test n'entame donc pas la perception d'infériorité des filles (elle ne semblent pas se rendre compte de cette réussite), sans que les instructions n'influencent dans un sens ou dans l'autre cette perception. Encore une fois nous soulignons que ce sentiment d'infériorité est en contradiction avec leur résistance à admettre le stéréotype.

Croyances sur la comparaison filles / garçons

Pour mesurer les croyances des sujets quant à la comparaison des scores attendus pour les deux sexes, la question suivante était posée aux participants : *Selon toi, comparativement aux garçons, les filles réussissent ce test : (1) beaucoup moins bien / ... / (5) beaucoup mieux.*

Échantillon Total. Si l'on demande aux élèves de comparer la performance présumée des filles et des garçons en général à ce test, l'opinion exprimée est que les filles sont sensées réussir légèrement mieux que les garçons¹⁵ ($M = 3.18$, $SD = .88$). L'ANOVA montre un effet de sexe, $F(1,871) = 80.58$, $p < .001$; $\eta_p^2 = .085$. Les filles estiment que leur groupe de sexe réussit mieux le test de RM ($M = 3.44$, $SD = .77$) que les garçons, alors que ceux-ci sont moins enclins à accepter l'idée d'une différence de sexe ($M = 2.92$, $SD = .92$) — ou tout au moins, l'estiment-ils très légèrement à leur avantage.

14. Un test-t pour un échantillon indique que la moyenne de l'auto-évaluation comparative est significativement supérieure à 3, $t(303) = 2.11$, $p < .05$.

15. Un test-t pour un échantillon indique que la moyenne de la comparaison filles / garçons est significativement supérieure à 3, $t(876) = 6.11$, $p < .001$.

Les différentes instructions n'ont eu aucune influence sur ces réponses, $F(2,871) < 1$, et de même l'interaction $Sexe \times Instruction$, $F < 1$, est non significative. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 72.24$, $p < .001$; $\eta^2 = .077$.

Échantillon Réduit. Nous observons dans l'échantillon sélectif que ces sujets ne jugent pas la performance comparée des filles et des garçons différente¹⁶ ($M = 3.09$, $SD = .83$). L'ANOVA montre un effet de sexe, $F(1,298) = 24.89$, $p < .001$; $\eta_p^2 = .077$, les filles affirmant toujours la supériorité de leur groupe de sexe ($M=3.41$, $SD = .65$), par rapport aux garçons ($M = 2.93$, $SD = .86$). Les effets des instructions, $F(2,298) < 1$, et de l'interaction $Sexe \times Instruction$, $F(2,298) < 1$, ne sont pas significatifs. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 8.87$, $p < .005$; $\eta^2 = .06$.

Croyances sur la comparaison jeunes / âgés

Pour mesurer les croyances des sujets quant à la comparaison des scores attendus entre des sujets jeunes et âgés, la question suivante était posée aux participants : *Selon toi, comparativement aux personnes âgées, les jeunes réussissent ce test : (1) beaucoup moins bien / ... / (5) beaucoup mieux.*

Échantillon Total. Si l'on demande aux élèves de comparer la performance présumée de sujets jeunes vs. âgés à ce test, l'opinion exprimée est que les jeunes sont sensées mieux réussir que les personnes âgées¹⁷ ($M = 3.43$, $SD = .99$). L'ANOVA montre un effet de sexe, $F(1,871) = 9.09$, $p < .005$; $\eta_p^2 = .010$. Les filles estiment que l'avantage des

16. Un test-t pour un échantillon indique que la moyenne de la comparaison filles / garçons n'est pas significativement différente de 3, $t(303) = 1.88$, *n.s.*

17. Un test-t pour un échantillon indique que la moyenne de la comparaison filles / garçons est significativement supérieure à 3, $t(876) = 12.91$, $p < .001$.

jeunes sur les sujets plus âgés est moins important ($M = 3.34$, $SD = .92$) que ce que ne l'estiment les garçons ($M = 3.53$, $SD = 1.06$), même si pour elles l'avantage des jeunes reste notable. Les différentes instructions n'ont eu aucune influence sur ces réponses, $F(2,871) < 1$, et l'interaction $Sexe \times Instruction$, $F(2,871) = 1.14$, est également non significative. L'effet de sexe est également observé suite à l'introduction du score au MRT en covarié, $F(1,870) = 8.87$, $p < .005$; $\eta^2 = .010$.

Échantillon Réduit. Nous observons dans l'échantillon sélectif que ces sujets jugent également dans l'ensemble que la performance des sujets jeunes devrait être supérieure à celle des sujets âgés¹⁸ ($M = 3.45$, $SD = 1.01$). L'ANOVA ne montre aucun effet de sexe, $F < 1$, filles ($M=3.37$, $SD = .91$) et garçons ($M=3.49$, $SD = 1.06$). Les effets des instructions, $F(2,298) = 1.23$, et de l'interaction $Sexe \times Instruction$, $F(2,298) = 1.07$, ne sont pas significatifs. Les effets ne sont également pas significatifs suite à l'introduction du score au MRT en covarié.

Au final, pour résumer, les filles semblent rejeter le stéréotype, tant du point de vue de sa diffusion dans la société que de leur adhésion personnelle. Surtout, elles déclarent même croire en une supériorité des femmes sur les hommes au test de RM, ce qui de fait constitue une croyance contre-stéréotypique. Dans le même temps, elles jugent ce test plus difficile que ne le font les garçons, et surtout lorsqu'il leur est demandé de juger leur propre performance et de se comparer aux autres élèves, elles déclarent individuellement moins bien réussir le test et se jugent défavorablement par rapport aux autres élèves. En bref, les reports verbaux traduisent l'influence du stéréotype négatif, surtout dans les questions où le stéréotype n'est pas explicitement évoqué.

18. Un test-t pour un échantillon indique que la moyenne de la comparaison filles / garçons est significativement supérieure à 3, $t(303) = 7.70$, $p < .001$.

3.4 Discussion

Les résultats de cette étude montrent clairement qu'un facteur social — un stéréotype négatif dotant les femmes d'une réputation d'incompétence dans le domaine visuo-spatial —, participe précisément à la régulation de l'expression de cette compétence chez les filles dès l'adolescence. Cette interférence peut donc expliquer, au moins pour partie, la différence de sexe en faveur des hommes classiquement observée en RM. En effet, si cette différence observée en RM était due à un facteur autre que l'interférence du stéréotype, ou du moins à un facteur indépendant du contexte social comme le serait par exemple un facteur biologique, alors cette différence devrait être imperméable à la manière dont la tâche est présentée, notamment par une simple phrase falsifiant le stéréotype. Or nous observons que cette falsification, qui contrarie les effets délétères du stéréotype interférent, conduit à l'annulation de la différence de sexe, notamment chez les élèves les plus performants en RM (les 30% obtenant les meilleurs scores). Or ceux-là constituent à la fois la population où les effets de sexe sont les plus massifs, mais aussi celle où la sensibilité au stéréotype est théoriquement la plus importante. Un fait notoire est que cette intervention falsifiante a conduit à une augmentation de la proportion de filles présentes dans cet segment de l'échantillon obtenant les meilleurs résultats en RM.

Ainsi, un des apports notables de notre étude est que nous avons choisi de tester un échantillon de très grande taille ($N=877$), ce qui nous a justement permis de quantifier l'évolution de l'effet du stéréotype le long de la distribution des performances en RM. Cela nous a permis d'isoler les individus pour lesquels cet effet devient critique : les filles susceptibles d'obtenir les meilleurs scores en RM. Notre contribution est particulièrement éclairante au regard du débat assez vif concernant les conséquences pratiques du phénomène en dehors du laboratoire. En effet, si le phénomène de menace du stéréotype a été largement répliqué dans de nombreuses situations évaluatives auprès de diverses populations stigmatisées (Désert et coll., 2002 ; Nguyen & Ryan, 2008 ; Wheeler & Petty, 2001), la taille de cet effet et sa portée dans des conditions écologiques sont encore peu rensei-

gnées, hormis dans la récente méta-analyse de Nguyen et Ryan (2008). Nous montrons ici que dans un cadre écologique (classe entière mixte), proportionnellement, le nombre de garçons présents dans les plus hauts niveaux de performance peut être plus de trois fois supérieur au nombre de filles au même niveau lorsque le stéréotype est activé, alors que la falsification de ce stéréotype permet de diviser ce facteur par deux. En outre, l'analyse des tailles d'effet dans cette partie de l'échantillon montre que cette falsification, en annulant les phénomènes de *stereotype threat* pour les filles, et de *stereotype lift* pour les garçons, peut même conduire les filles à surpasser les garçons ($d=-.25$).

Or, du point de vue de la recherche sur les différences de sexe, si les répliques du phénomène d'interférence du stéréotype dans le domaine des mathématiques a attiré l'attention, plusieurs auteurs de premier plan (p. ex. Halpern et coll., 2007), critiquent l'absence de données qui permettraient de cerner quelle part de la différence de sexe les effets du stéréotype pourraient expliquer, notamment dans des conditions écologiques. La controverse à ce sujet est très vive dans la littérature (c.f. l'échange entre Danaher & Crandall, 2008 ; Stricker & Ward, 2004, 2008). Or nous apportons ici des éléments d'importance pour clarifier ce débat. Nous montrons que si globalement — dans une population peu sélective — l'effet du *stereotype threat* sur la performance en RM n'est pas perceptible, ou est au moins statistiquement négligeable par rapport à l'effet de sexe, il devient critique lorsque l'on se focalise sur les filles et les garçons les plus doués en RM. Donc prendre la mesure de l'effet du stéréotype nécessite d'être vigilant quant aux caractéristiques de la population testée, car de manière pratiquement contre-intuitive, le sexe n'est pas le seul facteur déterminant dans ce cas, la sensibilité des individus aux effets du stéréotype dépend aussi de leur propre capacité. On peut se représenter ce phénomène comme une boucle de rétro-action négative : plus une fille possède des capacités en RM, et plus elle s'avère sensible aux effets négatifs du stéréotype qui en retour modèrent d'autant plus l'expression des capacités en question.

Un autre point très important que nous mettons à jour est que les effets du stéréotype ne sont pas seulement localisés dans la population testée, mais aussi temporellement. Afin d'être certain que le phénomène mis à jour par nos manipulation relevait bien de l'intervention d'un stéréotype, nous avons cherché à intégrer les développements théoriques récents concernant les différentes dimensions observables d'une telle interférence. Une description très précise du phénomène nous a donc permis de resserrer les critères qui signent sa survenue. La théorie du *stereotype threat* est le seul modèle qui prédit une baisse des performances des filles au fur et à mesure qu'elles avancent dans le test. En d'autres termes, la théorie prévoit que l'interférence attachée à l'intervention du stéréotype soit localisée dans la seconde section du test, une fois que l'accumulation de pensées négatives constitue une charge cognitive suffisante pour entamer la performance des filles les plus douées en RM. C'est exactement ce que nous observons. Nos résultats confirment non seulement les hypothèses avancées par Cadinu et coll. (2005), mais montrent aussi combien il convient d'être attentif à la dimension temporelle, souvent négligée, lorsque l'on cherche à montrer ladite interférence.

Le phénomène dont nous rendons compte est donc très dépendant du niveau de la population testée et de la phase (temporelle) de test considérée. Ceci, outre le caractère ténu des inductions, rend le *stereotype threat* difficile à mettre à jour, et peut donc expliquer le peu d'homogénéité des résultats de la littérature concernant les liens éventuels entre stereotype et RM. La subtilité des observables et des facteurs qui permettent d'observer le phénomène d'interférence lui confère à certains égards une coloration paradoxale, à plus d'un titre. En effet, si l'on prend en considération la sensibilité des inductions et du contexte expérimental, le phénomène de *stereotype threat* peut sembler versatile, sinon fragile : mais pourtant une simple phrase, falsifiant spécifiquement un stéréotype, peut annuler une différence de sexe par ailleurs connue pour être robuste. De plus, ce phénomène est particulièrement critique auprès de la population que l'on pourrait croire *a priori* immunisée contre les effets du stéréotype : les collégiennes réalisant les meilleures

performances en RM. Ces adolescentes affichant un caractère contre-stéréotypiques sont au contraire les plus sensibles à l'effet délétère du stéréotype.

Ces observations nous conduisent à noter une autre contradiction, qui apparaît si l'on se penche sur les déclarations des participants. D'un côté nous constatons un effet clairement protecteur de la falsification du stéréotype, réalisée à l'aide d'une phrase qui fait explicitement référence au stéréotype tout en le contredisant. Pourtant, nous constatons aussi, *via* les questionnaires, que les filles sont particulièrement peu enclines à adhérer au stéréotype et à le déclarer répandu dans la société. Et surtout, quelles que soient les instructions reçues, les filles affirment s'évaluer supérieures aux garçons au test de RM qu'elles viennent de passer et pour lequel on a pu mesurer un effet comportemental des différentes conditions. À l'inverse, elles déclarent le test de RM plus difficile que ne le considèrent les garçons et se jugent personnellement moins performantes que les autres participants.

Nous pouvons dégager deux pistes pour expliquer cette apparente contradiction. D'une part, nous pouvons y voir une certaine manifestation de la désirabilité sociale, témoignant qu'aujourd'hui il est socialement mieux accepté d'adopter un discours égalitariste — voire contre-stéréotypique — quant aux différences de sexes, quitte à ce que la majorité des participantes déclare que généralement, les filles réussissent un peu mieux les test spatiaux que les garçons. En outre, cette résistance face au stéréotype est certainement amplifiée par le fait que les filles défendent aussi la réputation de leur groupe social. Dans les faits, cette résistance s'exprime surtout lorsque les questions portent plus explicitement sur le stéréotype. Ainsi, nous constatons que cette résistance est plus forte lorsque l'on demande le degré d'adhésion au stéréotype que pour le degré de connaissance, cette seconde échelle étant déjà plus implicite que la première. La connexion entre le stéréotype et les questions qui portent sur la difficulté perçue et l'auto-évaluation au MRT est plus implicite encore, et précisément les réponses à ces questions vont dans le sens du stéréotype. Lorsque l'on demande de comparer explicitement la performance attendu des hommes et des femmes au

test, une résistance au stéréotype apparaît à nouveau. Par contraste, lorsque l'on évoque une comparaison entre des sujets jeunes et des sujets âgés, il n'y a aucune résistance d'exprimée, tant du côté des filles que des garçons, tous admettent que les personnes âgées sont moins performantes : c'est donc une résistance très localisée chez les filles au niveau du stéréotype de genre, qui les touche directement. Il est donc clair que les mesures auto-rapportées du stéréotype ne semblent être fiables que lorsqu'elles sont moins directes, plus implicites.

Une autre explication possible de cette contradiction dans les déclarations des filles pourrait être d'envisager un processus d'individuation. Les filles réfutent l'existence d'une mauvaise réputation qui stigmatiserait leur groupe social, mais portent sur elles-mêmes un jugement conforme au stéréotype (test perçu individuellement plus difficile, auto-évaluation comparative défavorable), quel que soit leur niveau de réussite effectif. Nous pourrions d'une certaine manière interpréter les réponses des filles en leur prêtant ce discours : « le stéréotype est faux, ce test est accessible aux filles, d'ailleurs — contrairement à ce que dit le stéréotype — elles le réussissent mieux que les hommes... *sauf moi*, je trouve d'ailleurs ce test difficile ». Cette seconde approche, si elle a le mérite d'ouvrir des perspectives sur les processus qui entrent en jeu lors de l'intervention d'un stéréotype (ici un processus d'individuation), reste une explication plus complexe que la précédente. Et surtout, elle n'explique pas pourquoi cette individuation s'observe dans toutes les conditions, alors que l'on pourrait s'attendre à ce que la falsification du stéréotype influence le positionnement des filles par rapport à leur groupe social : si l'individuation peut-être perçue comme un mécanisme de défense dans une situation de *stereotype threat*, la falsification qui neutralise le stéréotype devrait aussi agir sur l'individuation. Or ce n'est pas ce que l'on observe.

Quoi qu'il en soit, nos résultats sur les questionnaires sont à rapprocher des données de Huguet et Régner (2009) qui ont observé que les collégiennes qui réfutent le stéréotype ciblant les femmes en mathématiques ne sont pas plus protégées contre une baisse

de performance provoquée par une activation du stéréotype. Ces observations nous permettent de constater le rôle essentiel que joue la falsification explicite du stéréotype sur les performances des filles en RM, et ce, quelles que soient les connaissances et représentations qu'elles expriment explicitement à l'égard du stéréotype. Il est important ici de voir combien l'approche que nous adoptons se différencie de la plupart des modèles explicatifs des différences de sexe. Les modèles les plus consensuels aujourd'hui sont des modèles qui envisagent une explication des différences par une interaction sur le long terme entre l'expérience sociale de l'individu et ses aptitudes biologiquement acquises (Casey, Nuttall, & Pezaris, 1999 ; Halpern & Tan, 2001). L'originalité de notre approche porte essentiellement sur la manière dont sont envisagés les facteurs sociaux et cette interaction. Classiquement, les auteurs envisagent l'influence de ces facteurs sur le long terme, car la société offre par exemple moins d'opportunités pour les filles de cultiver leurs aptitudes spatiales, ceci créant un cercle vicieux les empêchant de compenser leur légère infériorité « biologique » dans ce domaine. Or ce type de modèle ne peut envisager une variation brutale de la performance des filles, comme celle que nous avons provoquée en modifiant simplement une phrase dans les instructions. Selon ces modèles, une variation de la performance n'advviendrait que sur le long terme, le temps par exemple que l'entraînement ou un renforcement social favorise l'établissement ou la modification de circuits cérébraux spécialisés et plus efficaces dans le traitement visuo-spatial. Pourtant, nous parvenons ici à manipuler la performance des filles en RM sur des délais extrêmement courts, c'est-à-dire les quelques minutes que dure la passation du MRT.

L'approche que nous avons adoptée apporte donc un regard original sur le rôle des facteurs sociaux dans la différence de sexe. Ces facteurs peuvent intervenir dans une situation de test, être manipulés expérimentalement — comme nous avons manipulé l'activation d'un stéréotype interférent —, et avoir un impact sur le comportement (performance) d'un individu dans des délais relativement brefs. De tels facteurs sociaux, qui sont des éléments du contexte dans lequel un individu agit, peut affecter la réalisation d'un traitement cognitif qui possède naturellement aussi un substrat biologique. L'infériorité des

filles en RM n'est plus selon nous à envisager comme un manque, un défaut d'aptitude inhérent à l'individu — que ce défaut résulte du lent processus de socialisation, ou soit inscrit biologiquement — mais plutôt comme une altération du traitement de l'information provoquée par une situation donnée. Ainsi la falsification permet à certaines collégiennes d'atteindre un niveau de performance comparable à celui des meilleurs collégiens, ce qui serait impossible si la capacité de réaliser de telles performances leur faisait foncièrement défaut. Cette différence de point de vue est essentielle pour appréhender les mécanismes qui peuvent expliquer les différences de sexe et envisager des interventions efficaces pour réduire cette différence, tant les facteurs à l'origine de cette différence peuvent être tenus et difficiles à saisir.

3.5 Conclusion

Nous avons montré que le paradigme de falsification de Spencer et coll. (1999) est parfaitement adapté pour mettre à jour l'existence d'un phénomène de régulation de la performance en RM par un stéréotype interférent. La falsification explicite du stéréotype se révèle une intervention efficace pour annuler la différence de sexe classiquement observée dans la frange de la population où cette différence est critique : les individus réalisant les meilleures performances. Nous montrons que la performance en RM n'est pas la pure et simple expression de la capacité d'un individu, mais le fruit d'une interaction entre cette capacité, l'identité sexuée de l'individu et un contexte social qui favorise ou non l'intervention (interférence) d'un stéréotype négatif ciblant les femmes. Nous avons également montré que la dynamique de la performance en RM doit donc être envisagée plus finement, notamment au cours du temps. L'efficacité du paradigme de falsification validé dans le cas du MRT, l'étude suivante reproduira ce paradigme dans un environnement écologique auprès d'étudiantes en licence de psychologie.

Chapitre 4

Interférence du stéréotype chez de jeunes adultes

4.1 Objectifs des études 2 et 3

Nous avons pu établir dans notre première étude que la falsification du stéréotype permettait de prévenir la sous-représentation des collégiennes dans les plus hauts niveaux de performance en RM, réduisant ainsi la différence de sexe là où elle est la plus visible. Dans le présent chapitre nous proposons deux études distinctes afin de confirmer ce résultat chez des étudiantes. L'étude 2 consiste en un questionnaire conçu pour déterminer si le MRT est spontanément perçu par les étudiants comme un test susceptible de mettre à jour une différence de sexe. En effet, l'existence d'attentes en rapport avec le stéréotype sur ce test serait un argument de poids en faveur de notre hypothèse que le test en lui-même active le stéréotype, ce qui conduit certaines femmes à une contre-performance. L'étude 3 vise à répliquer chez des étudiantes les résultats de notre première étude sur le MRT, où nous avons montré un effet positif de la falsification du stéréotype sur le score des collégiennes réalisant les meilleurs scores.

4.2 Étude 2 : attentes stéréotypiques concernant le MRT

L'étude 2 est conçue dans le but de montrer que le stéréotype est activé lors de l'administration du MRT par le test lui-même. Pour cela, nous proposons de mesurer le niveau des attentes des sujets sur la réussite de l'un ou l'autre sexe, lorsqu'ils sont confrontés au MRT. Cette étude se justifie par une question soulevée par nos résultats précédents. Si nous attribuons l'origine de la disparité des sexes constatée dans les hauts niveaux de performance à l'intervention d'un stéréotype, il nous faut prouver que la situation de test est propice à cette intervention. L'annulation de la différence de sexe consécutive à la falsification du stéréotype est un argument important en faveur de cette hypothèse. Néanmoins, d'autres indices peuvent renforcer notre interprétation, notamment en précisant quels sont les éléments de la situation de test qui provoquent l'activation du stéréotype. Or nous avons observé dans l'étude précédente que l'effet de sexe en condition contrôle, lorsque le test était présenté sans allusion explicite au stéréotype, était comparable à celui observé en condition d'activation explicite du stéréotype. Nous avons interprété ce résultat comme la preuve que, dans les conditions classiques de passation du test de RM, le stéréotype était activé.

Comment être sûr que le test, de par sa nature même, est suffisant pour activer le stéréotype en question ? Nous aurions pu introduire des mesures directes (p. ex. *via* des tâches de complétion de mot, cf. Steele & Aronson, 1995, ou de décision lexicale) de son activation immédiatement après les inductions, mais cela posait de nombreuses difficultés d'ordre méthodologique. En effet, nous avons proposé ce test dans un contexte écologique où le seul dispositif de mesure envisageable était un rapport verbal plus ou moins implicite. Des questions trop directes pouvaient conduire précisément à l'activation du stéréotype. L'activation mesurée ne serait alors pas due au test mais au questionnaire lui-même. En outre les rapports verbaux sont soumis à un biais de désirabilité sociale. Par exemple, nous

avons pu observer dans notre première étude que les filles étaient moins enclines que les garçons à déclarer adhérer au stéréotype et à le croire répandu. Pourtant, du point de vue comportemental, la falsification s'est avérée efficace chez elles pour prévenir la différence de sexe habituellement observée en leur défaveur, ce qui selon nous est difficilement explicable en dehors de l'hypothèse d'une neutralisation du stéréotype. Ceci confirme les résultats de (Huguet & Régner, 2009) qui montrent que des collégiennes niant le stéréotype de genre en mathématiques y sont néanmoins sensibles.

Une solution envisagée dans le présent chapitre consiste à vérifier les attentes spontanées des participants face au test de RM (au moins concernant les femmes). En effet, si l'existence d'un stéréotype en rapport avec les aptitudes spatiales est acceptée (Blakemore et coll., 2009 ; Halpern, 2000), peu d'études l'ont mis directement en évidence (Robert, 1990 ; Halpern, 2000). Il est bien connu, pour le sens commun, que « *les femmes n'ont pas le sens de l'orientation* », « *ne savent pas lire une carte routière* », ou « *sont mauvaises en géométrie* ». Nous avons donc de bonnes raisons de croire que le domaine visuo-spatial est bien reconnu comme un domaine masculin, même si évidemment, il n'existe pas vraiment de formulation explicite du stéréotype, par exemple « *les femmes sont mauvaises dans le domaine visuo-spatial* », comme on peut l'entendre au sujet des mathématiques.

Dans la littérature scientifique, la réussite en RM est également associée assez souvent à des compétences masculines (Hamilton, 1995 ; Signorella et coll., 1989) et à la pratique d'activités spécifiquement masculines (Alington et coll., 1992 ; Blakemore et coll., 2009 ; Casey, 1996b ; Newcombe et coll., 1983). Ces observations ne sont pas satisfaisantes pour expliquer la différence de sexe, du fait que les études corrélationnelles ne donnent aucune indication sur les relations causales. Mais peut-être les représentations qui découlent du constat de cette association avec des traits masculins sont elles suffisantes pour diffuser et surtout renforcer le stéréotype. Cette attente, dans la situation de test, se révélerait alors délétère pour la performance des femmes selon l'hypothèse d'une intervention du stéréotype.

Nous proposons donc de mesurer, via un questionnaire, s'il existe parmi les étudiant(e)s une telle croyance en la supériorité des hommes en RM. Si nous montrions que le test de RM en lui-même est associé à une attente de supériorité des hommes, cette observation serait un argument supplémentaire, avec l'efficacité de la falsification, pour établir le rôle du stéréotype dans l'expression de la différence de sexe en RM. Le questionnaire consistera en une série d'exemples de problèmes tirés de différents tests cognitifs (dont le MRT) pour lesquels les participants devront indiquer le degré de réussite qu'ils attendent pour différents groupes (femmes, hommes, jeunes, âgés, eux-mêmes, etc.).

4.2.1 Méthode

Participants

Nous avons recueilli les attentes de 62 étudiants (21 hommes ; 41 femmes) en première année de psychologie (L1) des universités de Provence (9 hommes ; 18 femmes) et de Nantes (12 hommes ; 23 femmes). Une femme a été écartée en raison de son âge (reprise d'étude), portant le nombre total de femmes à 40 et de participants à 61 (âge moyen = 20.55 ans). Le questionnaire a été soumis à des étudiants n'ayant pas encore reçus leurs premiers cours de psychologie différentielle, où ils auraient pu être sensibilisés à la problématique des différences de sexe dans le domaine cognitif, notamment visuo-spatial.

Procédure

Les participants répondaient individuellement au questionnaire, proposé en situation collective (séances de TD de psychologie cognitive ou de méthode expérimentale). Le questionnaire était proposé en début de séance, présenté comme un support du cours à venir. L'expérimentateur précisait bien qu'il était important de répondre non seulement sérieusement, mais « de la manière la plus spontanée possible, le but étant de recueillir la

première pensée qui vous vient à l'esprit. Il n'y a pas de bonne et de mauvaise réponse ». Le temps de réponse était libre (une quinzaine de minute). Les questionnaires distribués, les participants indiquaient, comme pour les tests de RM proposés dans les autres études, leur date de naissance et leur sexe sur la première page du questionnaire. Puis ils disposaient de 3 minutes pour lire l'énoncé et répondre à l'item donné en exemple. A l'issue de ce temps, l'expérimentateur répondait aux questions éventuelles, puis donnait le départ.

Questionnaire

Le questionnaire figure en annexe p. 238. Les problèmes présentés étaient à chaque fois un exemple tiré des test suivants :

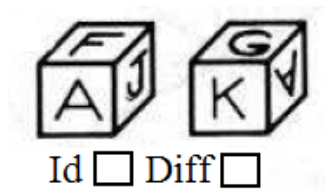
- fluence verbale (produire le plus de mots possibles commençant par une lettre donnée) ;
- rotation mentale (MRT, Vandenberg & Kuse, 1978) ;
- matrices de Raven (raisonnement, Raven, Raven, & Court, 1998) ;
- une liste de mots à retenir (mémoire) ;
- Identical Block Test (IBT — test visuo-spatial, (Stafford, 1961)) ;
- Water Level Test (WLT — test visuo-spatial, Vasta, Belongia, & Ribble, 1994, d'après Piaget & Inhelder, 1967).

Un exemple de page du questionnaire figure p. 121 (fig. 4.1). Pour chaque problème présenté la réponse correcte était indiquée, afin que les réponses des sujets ne soient pas orientées seulement par leur capacité à résoudre le problème donné en exemple. Les participants indiquaient sur des échelles en sept points leurs attentes sur la réussite de différentes groupes *cibles* au problème. Ces cibles correspondaient à différents groupes, comme des individus jeunes, âgées, d'un bon (ou faible) niveau en science, etc. Enfin, pour chaque type de problème, une dernière échelle concernait l'auto-évaluation du participant.

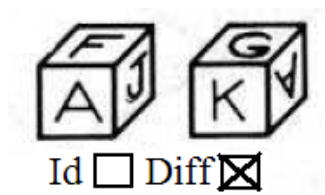
4.2 Étude 2 : attentes stéréotypiques concernant le MRT

Dans ce test, il faut indiquer si deux cubes posés côte à côte peuvent être identiques (dans ce cas il faut cocher la case « Id ») ou différents (dans ce cas il faut cocher la case « Diff »).

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :



Réponse :



Selon vous, sur ce test, le résultat des personnes jeunes est en général :

IB1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

IB2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

IB3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

IB4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

FIGURE 4.1 – Page tirée du questionnaire conçu pour mesurer les attentes des étudiants sur différents tests (ici IBT). Pour chaque problème donné en exemple, les participants indiquaient leurs attentes concernant la réussite de différentes populations cibles (ici seules quatre échelles sont données en exemple, voir l'annexe p. 238 pour la totalité des questions).

Notre objectif étant bien de mesurer les attentes concernant la différence de sexe en RM, seules les attentes concernant le problème de RM et ciblant les groupes de sexe étaient critiques pour notre problématique. Nous considérons les questions portant sur les autres problèmes et les autres groupes cibles (jeunes, âgés, etc.) comme des items de remplissage, visant à ne pas rendre trop abrupte la question de la comparaison des hommes et des femmes en RM. Néanmoins il est utile de montrer que, par contraste avec les autres types de problème, les attentes stéréotypées ne concernent que les problèmes de RM ou visuo-spatiaux. Dans la même idée, afin de limiter les problèmes de désirabilité sociale, les questions ne consistaient pas en des comparaisons directes entre les deux sexes, mais en des évaluations séparées de leur performance.

Enfin, dans le but de progresser dans l'interprétation de l'étude 1, nous introduisons de la même manière une échelle d'auto-évaluation. Nous avons alors observé que les filles s'auto-évaluaient moins positivement en RM que les garçons alors qu'elles manifestaient par ailleurs une résistance à adhérer ou même seulement à reconnaître l'existence du stéréotype. Si, dans le présent questionnaire, les étudiantes affichent la même résistance envers le stéréotype lorsqu'il est évoqué trop explicitement, cette mesure plus implicite permettra de confirmer notre observation précédente, et de montrer que cette auto-évaluation plus négative n'est pas dépendante de la passation du test, mais constitue une attente spontanée vis à vis du MRT.

Attentes

Nous nous attendions à déceler l'existence d'un stéréotype de sexe concernant les attentes en RM, les femmes devant être considérées comme moins performantes que les hommes. Au vu des résultats des questionnaires post-test de la précédente étude, il est envisageable néanmoins que les femmes repoussent l'idée d'une différence de sexe en général — donc rapportent des attentes identiques pour les performances des deux sexes — mais se considèrent individuellement comme moins capables (auto-évaluation) que les hommes.

Nous nous attendions aussi à ce que les personnes âgées soient considérées comme moins performantes que les sujets jeunes en RM, mais aussi dans la plupart des tâches, le déclin cognitif avec l'âge étant supposé général (concentration, mémoire, planification, etc.).

4.2.2 Résultats

Attentes concernant le MRT

Attentes pour les groupes de sexe. Nous avons procédé à une ANOVA mixte *Sexe* \times *Cible* avec ce dernier facteur en mesure répétée. Pour les problèmes de RM, nous observons un effet principal du sexe des participants, $F(1, 59) = 11.98, p < .001 ; \eta^2 = .17$, les femmes attendent dans l'ensemble de meilleurs résultats ($M = 4.90, SD = 1.08$) que les hommes ($M = 4.21, SD = .73$). Nous observons également un effet principal significatif du groupe ciblé (féminin vs. masculin), $F(1, 59) = 34.90, p < .001 ; \eta^2 = .37$, les attentes concernant les hommes étant globalement supérieures ($M = 5.03 ; SD = .86$) à celles concernant les femmes ($M=4.30, SD = 1.17$), conformément au stéréotype. Cet effet est essentiellement localisé chez hommes, comme le montre l'interaction significative $Sexe \times Cible$, $F(1, 59) = 16.11, p < .001 ; \eta^2 = .21$, (cf. fig. 4.2, p. 124). En effet, les femmes s'attendent à une légère supériorité des hommes ($M = 5.05, SD = .88$) sur les femmes ($M = 4.75, SD = .82$) alors que les hommes s'attendent à une réussite de leur groupe de sexe bien plus marquée ($M = 5.00, SD = .91$) par rapport aux femmes ($M = 3.43, SD = 1.21$).

Auto-évaluation. Globalement, les participants ont tendance à s'auto-évaluer positivement sur le MRT ($M = 4.61, SD = 1.01$), un test-t pour échantillon unique indique que la moyenne de l'auto-évaluation est significativement supérieure à 4, $t(60) = 4.72, p < .001$. Une ANOVA à un facteur montre néanmoins un effet significatif du sexe des participants sur leur auto-évaluation, $F(1, 59) = 5.26, p < .05 ; \eta^2 = .082$, les femmes

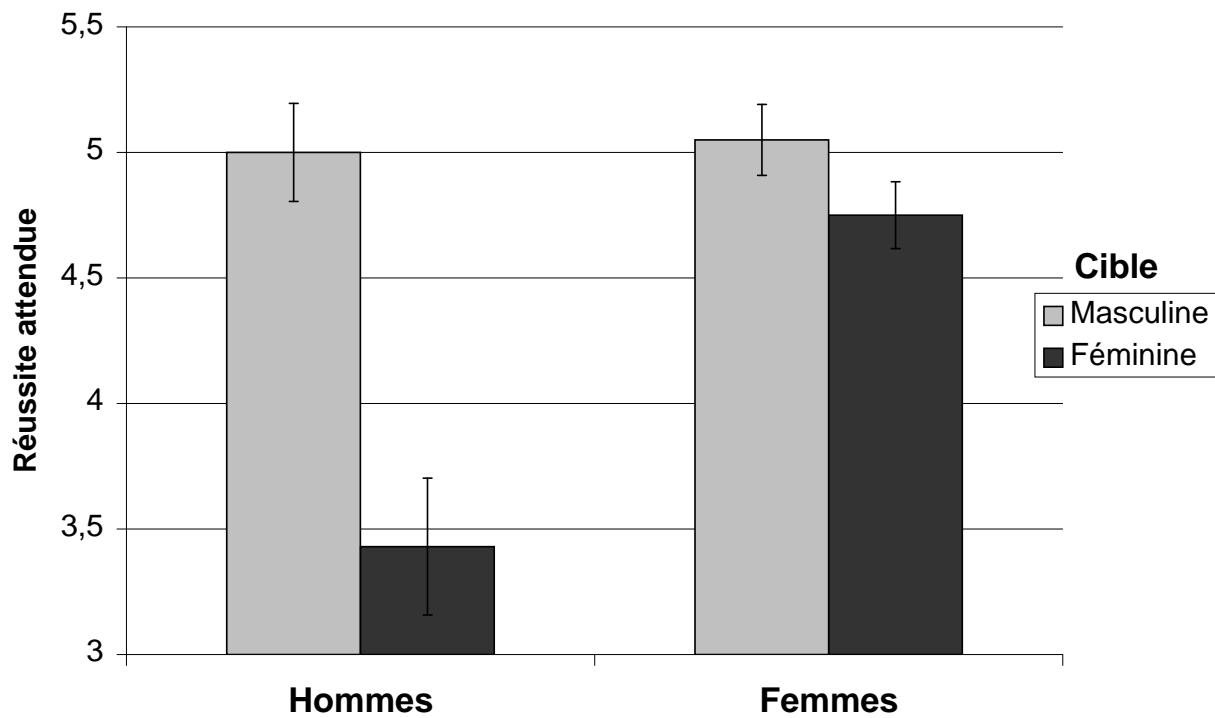


FIGURE 4.2 – Interaction *Sexe* \times *Cible* sur la réussite attendue au MRT (1 - Très mauvais / 7 - Très bonne réussite)

s'auto-évaluent un peu moins positivement ($M = 4.40$, $SD = .67$) que ne le font les hommes ($M = 5.00$, $SD = .96$). Cet effet du sexe confirme les résultats observés sur le questionnaire de la première étude. En effet, alors que les femmes semblent résister pour déclarer l'existence d'une différence entre les sexes lors d'une comparaison explicitement intersexe, elles s'auto-évaluent de manière moins favorable que les hommes.

Attentes concernant l'âge. Une ANOVA mixte montre un net effet principal de la cible pour le problème de RM, $F(1, 59) = 36.33$, $p < .001$; $\eta^2 = .38$, hommes et femmes considérant que les sujets jeunes réussissent mieux le test ($M = 4.68$, $SD = 1.12$) que les personnes âgées ($M = 3.43$, $SD = 1.07$). Aucun effet du sexe, $F(1, 59) = 1.19$; *n.s.*, et aucune interaction ($F < 1$) n'apparaissent.

Attentes pour les autres problèmes

Les autres problèmes pour lesquels des attentes étaient recueillies consistaient en deux tests appartenant au domaine verbal (fluence et mémorisation), deux tests visuo-spatiaux (IBT et WLT), et un test de raisonnement (matrices de Raven).

Attentes pour les groupes de sexe. Les ANOVA révèlent un effet principal de sexe pour trois types de test — Fluence verbale, $F(1, 59) = 8.74, p < .005 ; \eta^2 = .13$, Matrices de Raven, $F(1, 59) = 6.59, p < .005 ; \eta^2 = .10$ et IBT, $F(1, 59) = 5.64, p < .005 ; \eta^2 = .089$ — dont seul le dernier relève du domaine visuo-spatial. Pour ces trois tests, les femmes ont tendance à rapporter des attentes supérieures à celles des hommes. Mais surtout, nous observons un effet principal de la cible significatif pour d'autres problèmes, à chaque fois dans des directions en accord avec le sens prévu par les stéréotypes de genre (cf. tableau 4.1). Ainsi, pour le WLT — test visuo-spatial — les hommes sont crédités d'une réussite significativement supérieure à celle des femmes, $F(1, 59) = 4.03, p < .05 ; \eta^2 = .064$. Pour les tests portant sur du matériel verbal, l'effet du groupe ciblé est aussi significatif — Fluence verbale, $F(1, 59) = 9.87, p < .005 ; \eta^2 = .14$ et en mémorisation de mots, $F(1 ; 59)=5,15 ; p<,05 ; \eta^2 = .080$ —, les meilleures performances sont attendues pour les femmes, conformément au stéréotype qui leur confère une réputation de supériorité dans le domaine verbal (cf. chapitre 1 p. 21). Cependant, nous constatons que les différences de sexe attendues sur ces tests sont de taille plus réduite que celles concernant la RM. Les analyses ne mettent à jour aucune interaction, contrairement à nos observations pour le test de RM.

Auto-évaluation. Les participants des deux sexes s'évaluent positivement à l'ensemble des tests (les auto-évaluations sont toutes supérieures à 4, $p < .001$), et l'ANOVA ne révèle aucun effet de sexe (tous les $F < 1$ sauf pour la fluence verbale où $F(1, 59) = 1.72, n.s.$).

TABLE 4.1 – Attentes concernant la réussite des hommes et des femmes à différents tests.

Cible		Type de test			
		WLT	Fluence verbale	Mémorisation de mots	MRT
Hommes	<i>M</i>	5.16	4.89	4.69	5.03
	<i>SD</i>	.94	1.02	.86	.86
Femmes	<i>M</i>	5.00	5.34	4.93	4.30
	<i>SD</i>	1.09	.94	.86	1.17

Attentes concernant l'âge. Les ANOVA mixtes révèlent pour plusieurs tests un effet du groupe ciblé, avec une différence attendue dans le sens d'une moindre réussite des personnes âgées par rapport aux jeunes. Cet effet est observé pour le test des Matrices de Raven, $F(1, 59) = 21.50$, $p < .001$; $\eta^2 = .27$, pour la mémorisation de mots, $F(1, 59) = 90.58$, $p < .001$; $\eta^2 = .61$, et pour l'IBT, $F(1, 59) = 12.75$, $p < .001$; $\eta^2 = .18$. Nous remarquons que la taille d'effet la plus importante est observée pour le test de mémoire, conformément au stéréotype sur le vieillissement cognitif. Les attentes pour les deux autres tests, relativement difficiles et nécessitant une certaine fluidité mentale, sont également compatible avec un stéréotype au sujet du vieillissement cognitif. Par ailleurs, l'ANOVA révèle un effet du sexe des participants significatif pour le test de fluence verbale, $F(1, 59) = 5.67$, $p < .005$; $\eta^2 = .088$, les femmes jugeant le test globalement un peu mieux réussi ($M = 4.79$, $SD = .85$) que ne le font les hommes ($M = 4.24$, $SD = .86$).

4.2.3 Discussion

Les résultats montrent qu'il existe bien des attentes différentes quant à la performances en RM d'individus appartenant à des groupes sociaux différents. La plupart des auteurs qui se sont penchés sur cette question montrent seulement un lien entre la réussite en RM et la possession de traits typiquement masculins, comme la préférence pour des activités telles que le sport ou la mécanique (Signorella & Jamison, 1986). Comme il s'agit d'activités stéréotypiquement masculines, les auteurs en ont conclu que les participants

confrontés au test de RM perçoivent probablement ce test comme typiquement masculin¹. Nos résultats montrent clairement que la perception de différents tests cognitifs concorde bien avec les stéréotypes de genre. Les tests associés au domaine verbal, typiquement féminins, sont perçus mieux réussis par les femmes que par les hommes, tandis que les attentes concernant les tests visuo-spatiaux (MRT, WLT) vont dans le sens — comme le stéréotype le suggère — d’une supériorité des hommes. Le WLT est un test relativement facile, pour lequel l’effet de sexe rapporté dans la littérature est de taille moyenne, alors que le MRT est un test difficile, pour lequel l’effet de sexe rapporté est grand (cf. chapitre 1 p. 25). Or les participants que nous avons testés s’attendent précisément à une différence bien moins importante au WLT qu’au MRT. Les attentes spontanées de sujets naïfs confrontés pour la première fois à ces tests correspondent donc aux différences de sexe observées dans la littérature, ce qui constitue un argument en faveur du rôle que pourrait jouer ces attentes en question dans l’expression de ces différences.

Par ailleurs les participants n’attendaient aucune différence de sexe pour les Matrices de Raven, en dépit de quelques résultats qui semblent indiquer l’existence d’un stéréotype de genre dans le domaine du raisonnement (Davies, Spencer, Quinn, & Gerhardstein, 2002 ; Heilman, Block, & Martell, 1995). Les participants n’attendaient pas davantage de différence de sexe à l’IBT, bien qu’il s’agisse d’un test visuo-spatial, et surtout un test qui est connu pour montrer des différences de sexe de même magnitude que le MRT (Voyer et coll., 1995). L’IBT ici n’est pas perçu comme un test très discriminant entre les sexes. Il est possible que cela soit dû au fait que l’IBT, juxtaposé au MRT, donne l’impression d’être beaucoup plus simple. Le MRT porte sur des objets complexes composés de multiples cubes, alors que les problèmes de l’IBT portent sur la reconnaissance d’un cube unique dont les faces sont identifiées par des symboles simples. Dans le même ordre d’idée, on peut noter que la différence attendue au WLT, un autre test assez facile, est petite —

1. A notre connaissance seule Robert (1990), à la suite de McMahan (1982), avait tenté plus directement d’établir si un test visuo-spatial — le WLT — était perçu comme typiquement mieux réussi par les hommes, pour conclure que la perception de cette tâche était neutre. Pour ce même test, nous montrons ici qu’une différence d’attente existe en faveur des hommes, donc dans le sens du stéréotype, bien que cette différence soit très réduite (la plus petite de tous les tests pour lesquels une différence d’attente existe).

Robert (1990) n'avait d'ailleurs pas non plus pu montrer une différence dans les attentes concernant la réussite des deux sexes au WLT. Par contre nous voyons que le MRT est lui associé à des attentes clairement stéréotypiques. Le MRT serait donc perçu de manière relativement singulière, comme une tâche spatiale très difficile, et donc mettant très nettement en évidence les différences de sexe.

Du reste, le MRT est le seul test pour lequel nous avons observé une interaction *Sexe* \times *Cible*. Par rapport aux hommes, les femmes s'attendent à une différence de sexe bien moins marquée à ce test. Ce qui éventuellement corrobore leur résistance à exprimer des attentes clairement compatibles avec le stéréotype. À l'inverse, lorsque le stéréotype leur est favorable (test verbaux) leurs résistances à exprimer des attentes stéréotypiques disparaissent. Ces résultats sont compatibles avec ceux concernant les auto-évaluations. La seule tâche pour laquelle les femmes s'auto-évaluent significativement inférieures aux hommes est précisément le MRT. Elles déclarent, nous l'avons vu précédemment, que les femmes dans leur ensemble devraient obtenir des résultats juste inférieurs à ceux des hommes en RM, alors qu'elles-mêmes s'estiment individuellement (auto-évaluation) nettement inférieures. Cela rejoint nos précédentes observations en référence à l'étude 1, où les collégiennes s'estimaient moins performantes en RM mais indiquaient ne pas croire en une différence de sexe à leur désavantage. Nous avons conclu que les filles résistaient à une mesure trop explicite du stéréotype, ce qui semble donc être le cas ici aussi.

Enfin, un autre stéréotype, ciblant l'âge, est plus franchement mis en évidence. Nous remarquons que les tests où l'âge du groupe cible influe sur les attentes concernent des domaines cognitifs visés par le stéréotype sur l'âge : mémoire, raisonnement, manipulation active de l'information spatiale, alors que les tests épargnés par ce jugement mettent en jeu des connaissances acquises (vocabulaire, horizontalité d'un liquide). Ce résultat est un élément en faveur d'une manipulation alternative à la falsification explicite du stéréotype que nous envisageons dans l'étude 3. En effet, si les femmes s'attendent bien à obtenir des résultats supérieurs au MRT par rapport aux personnes âgées, alors nous pouvons

envisager que les inciter à cette comparaison au moment où elles s'engagent dans le test aura un effet positif sur leur performance. En activant une autre dimension qui leur est plus favorable, cette comparaison préviendrait l'activation du stéréotype de genre.

4.3 Étude 3 : Intervention alternative à la falsification

Dans cette étude 3, notre objectif principal était de montrer, chez des étudiantes, la reproductibilité de l'effet positif de la falsification explicite du stéréotype, précédemment mis en évidence chez les collégiennes lors de l'étude 1. Nous appuyant sur les résultats de l'étude 2, il s'agissait aussi de montrer que cet effet positif peut également être obtenu à l'aide d'une procédure focalisant l'attention sur une comparaison sociale avantageuse (des jeunes avec les âgés). En effet certaines données indiquent qu'une comparaison avec les membres d'un exogroupe sont menaçantes pour le soi lorsque le domaine de la comparaison renvoie à un stéréotype négatif, comme c'est le cas lorsque des jeunes femmes sont amenées à se comparer avec des hommes dans le domaine de la RM. Par contre, la comparaison avec les membres du même groupe social ou avec les membres d'un exogroupe sans relation avec un stéréotype est moins menaçante (Martinot & Redersdorff, 2003 ; Redersdorff & Martinot, 2003). Ici, en suggérant une comparaison avec un exogroupe ciblé par un stéréotype négatif à l'avantage des jeunes femmes, nous envisageons même un effet positif de cette comparaison sur la performance des femmes en RM. Nous porterons une attention particulière à la dynamique temporelle de l'effet du stéréotype, localisé dans la seconde section du test. C'est donc dans cette dernière section que les effets positifs de la falsification devraient être les plus sensibles.

4.3.1 Méthode

Participants

Notre échantillon était composé de 278 femmes (âge moyen = 19.48). Nous avons initialement testé 381 étudiants (39 hommes ; 314 femmes ; 28 non renseignés) en première et en seconde année de psychologie (L1 / L2) de l'université de Nantes. Le test était proposé avant les premiers cours de psychologie différentielle, dont le programme portait sur la différence de sexe dans les activités visuo-spatiales. Les 103 protocoles rejetés l'ont été pour les raisons suivantes : 39 protocoles recueillis auprès d'hommes, trop peu nombreux pour constituer des groupes comparables aux groupes de femmes, 9 protocoles de femmes dont l'âge était supérieur à 25 ans et enfin 55 protocoles incomplets (sexe ou âge non-renseigné, test ou questionnaire incomplet, retard, etc.) ou de sujets déjà confrontés au test par le passé.

Procédure

Le plan expérimental était du type $3(Instructions) \times 2(Section)$. Les étudiantes étaient réparties ainsi entre les différentes conditions : 73 dans la condition standard (contrôle), 112 en falsification et enfin 93 participantes dans la condition où les instructions suggéraient une comparaison avec des personnes âgées.

Hormis la présence d'une condition différente (comparaison), la procédure était rigoureusement la même que celle mise en œuvre lors de l'étude 1 (cf. chapitre 3, p. 85). L'expérimentateur, de sexe masculin, expliquait le déroulement de l'étude en distribuant les différentes parties du test face contre table. Pour décourager les étudiantes de copier les unes sur les autres, des tests étaient distribués en alternance, et bien qu'identiques, portaient des intitulés différents (« test A » ou « test B ») imprimés sur des feuilles de couleurs également différentes (blanche ou jaune clair).

Dans un premier temps, les participantes devaient indiquer sur la première page leur date de naissance et leur sexe, et enfin lire les instructions du MRT. La dernière phrase des instructions du MRT constituait l'induction correspondant à la condition expérimentale. Les inductions pour les conditions *standard* et *falsification* étaient identiques à celles employées dans les collèges (cf. p. 86). Pour la nouvelle condition *comparaison*, l'induction était la suivante : « *Pour information, cette session de test consiste en une simple condition contrôle auprès d'une population jeune, dans une étude portant sur la performance des personnes âgées (>70 ans) à ce test.* »

Ensuite l'expérimentateur répondait aux éventuelles questions, puis donnait trois minutes aux participantes pour répondre à la première partie du MRT. Après une pose d'une minute, les étudiantes répondaient à la seconde partie du MRT toujours selon les mêmes modalités. À l'issue de cette seconde partie, un questionnaire (cf. annexe p. 252) était soumis aux participantes. Dans ce dernier, nous avons choisi de mesurer l'adhésion au stéréotype, déterminée par quatre questions formulées de manière plus directes que dans l'étude 1. Pour ces questions nous avons choisi une échelle en 6 points, afin de forcer les sujets à se positionner dans le sens ou à l'opposé du stéréotype. Les autres questions consistaient en des échelles à cinq points pour mesurer l'auto-évaluation et plusieurs questions d'auto-évaluation comparative (par rapport aux autres participantes, aux hommes, aux femmes et enfin aux personnes âgées). Deux échelles distinctes étaient destinées à mesurer les attentes des participantes par rapport à la réussite comparée d'individus de sexe et d'âge différents, et enfin une échelle concernait la difficulté perçue du test.

Attentes

Nous nous attendions dans un premier temps à observer dans les conditions de *falsification* et de *comparaison* (avec les personnes âgées) un plus grand nombre de femmes dans les hauts niveaux de performance par rapport à la condition *standard*, et ce notamment dans la seconde section du test. Nous nous attendions par ailleurs à une interaction

Instruction \times *Section*, caractéristique de la dynamique du *stereotype threat*. Plus spécifiquement, nous attendions un score identique entre la première et la seconde section du test dans les conditions *falsification* et *comparaison*, et une baisse de la performance entre les deux sections dans la condition *standard*, sensée activer implicitement le stéréotype.

4.3.2 Résultats

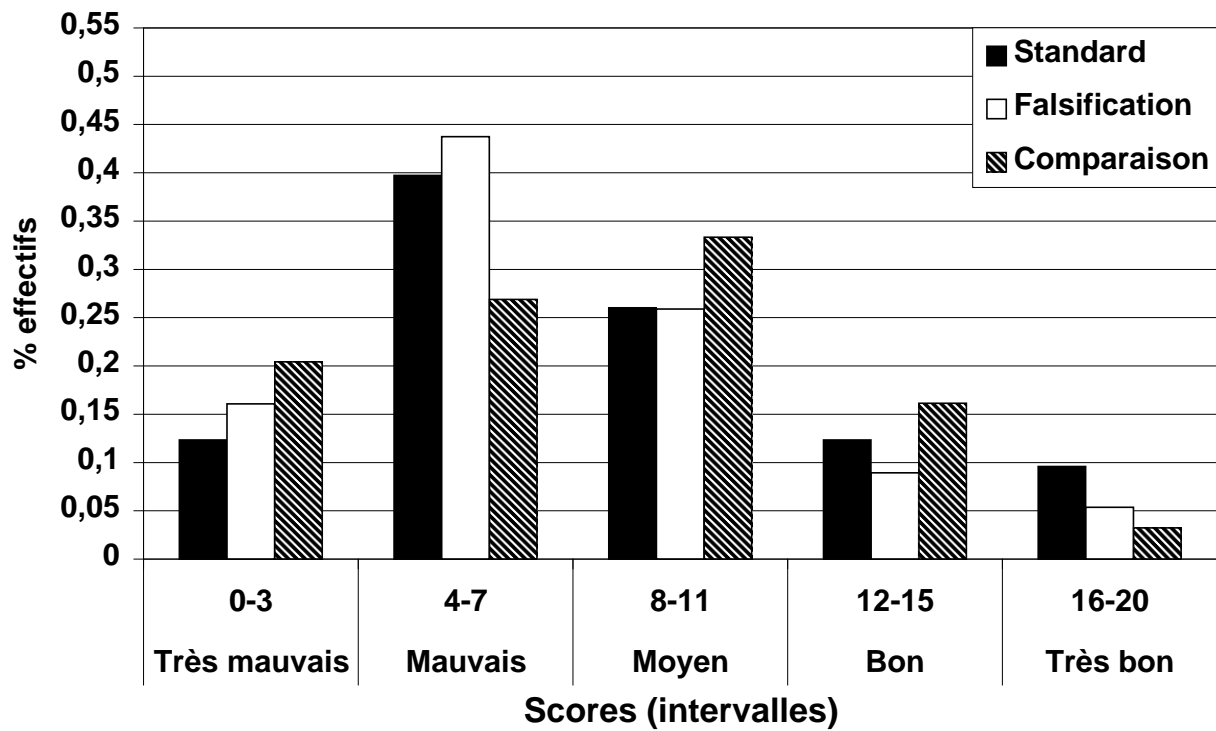
MRT

Niveaux de performance. Pour comparer les distributions de performance observées dans chaque condition, nous ne pouvions pas utiliser la même technique d'analyse que lors de l'étude 1, du fait que seules des femmes ont été testées, ce qui rend impossible le calcul des sexe-ratios. Nous avons donc choisi de diviser l'échelle des scores en cinq tranches ou niveaux de performance, soit : *très mauvais*, *mauvais*, *moyen*, *bon* et *très bon*. Le niveau *très mauvais* regroupait les scores de 0 à 3², le niveau suivant (*mauvais*) regroupait les scores de 4 à 7, etc., jusqu'au niveau *très bon* (score de 16 à 20). Pour chaque condition, nous avons calculé le rapport entre le nombre de participantes dans chaque niveau de performance et le nombre total de participantes dans la condition considérée (cf. fig. 4.3 p. 134).

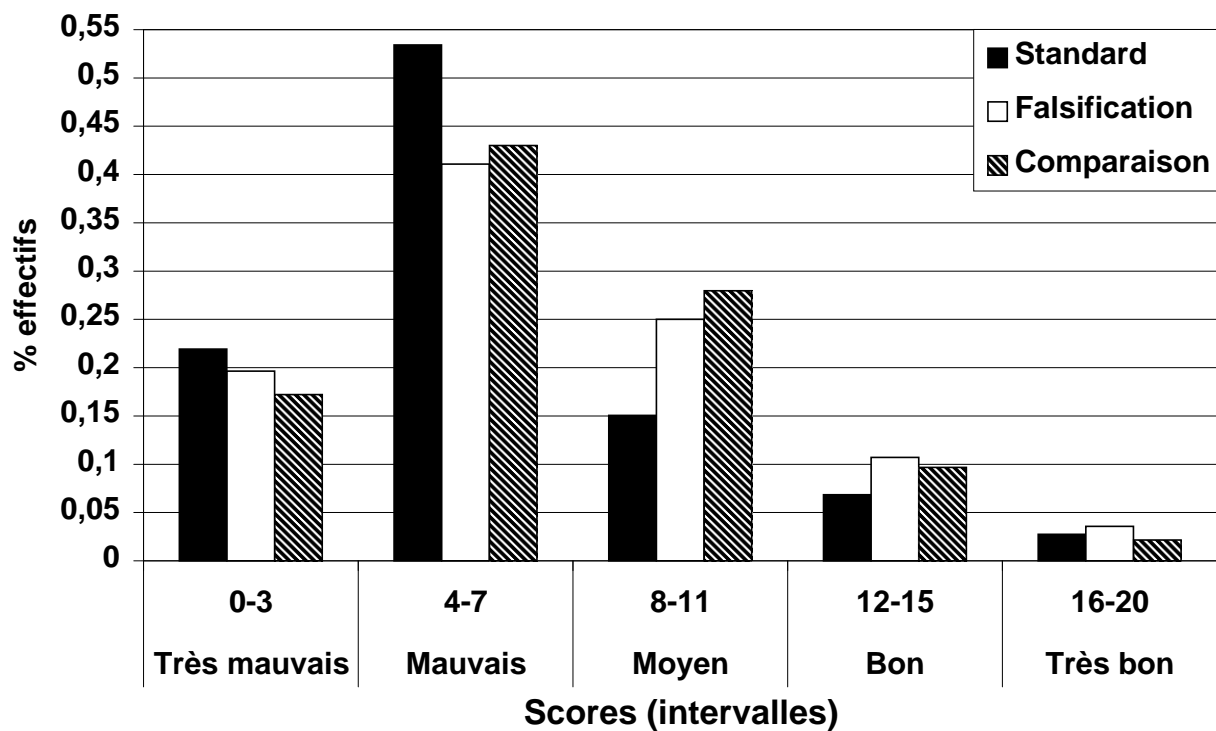
Dans la première partie du test, on ne distingue dans l'ensemble pas d'écart important (moins de 10% en général) entre les effectifs de chaque condition aux différents niveaux de performance, ce qui est conforme à nos hypothèses. Cependant, dans la seconde section du test, nous n'observons pas dans les hauts scores la différence attendue entre les distributions des différents groupes. En fait, peu d'étudiantes atteignent ces hauts scores et donc les effectifs des groupes comparés à ces niveaux de performance sont très faibles. Ainsi, pour la seconde section du test, seulement 10% des participantes réalisent dans la condi-

2. Comme une dimension caractéristique du phénomène est sa dynamique temporelle, nous nous sommes focalisé sur la différence entre les distributions de performance entre la première et la seconde section du MRT, pour chaque partie le score maximum était de 20.

tion *standard* un score de *bon* à *très bon*, 14% en *falsification*, et 12% dans la condition de *comparaison*. Même si la taille de notre échantillon total est conséquente ($N = 278$), les mauvaises performances des participantes ne nous permettent pas d’isoler objectivement suffisamment de sujets pour réaliser des analyses fiables dans les hauts niveaux de performance. Nous devons donc renoncer à ces analyses. Tout au plus peut-on constater entre les différentes conditions un léger décalage des distributions de performance dans le sens attendu : les chiffres donnés précédemment indiquent que les étudiantes en condition de *falsification* sont légèrement plus nombreuses à obtenir un score élevé, tandis qu’il apparaît que les étudiantes en condition *standard* sont majoritaires dans les niveaux *mauvais* et *très mauvais* (cf. fig. 4.3). Ceci suggère que si l’effet existe, il est de très petite amplitude. Aussi, si nous devons renoncer à une analyse en fonction des niveaux de performance, l’analyse de la dynamique du phénomène, par la comparaison des résultats entre la première et la seconde section du test, n’en demeure pas moins informative.



(a) Première partie du MRT



(b) Seconde partie du MRT

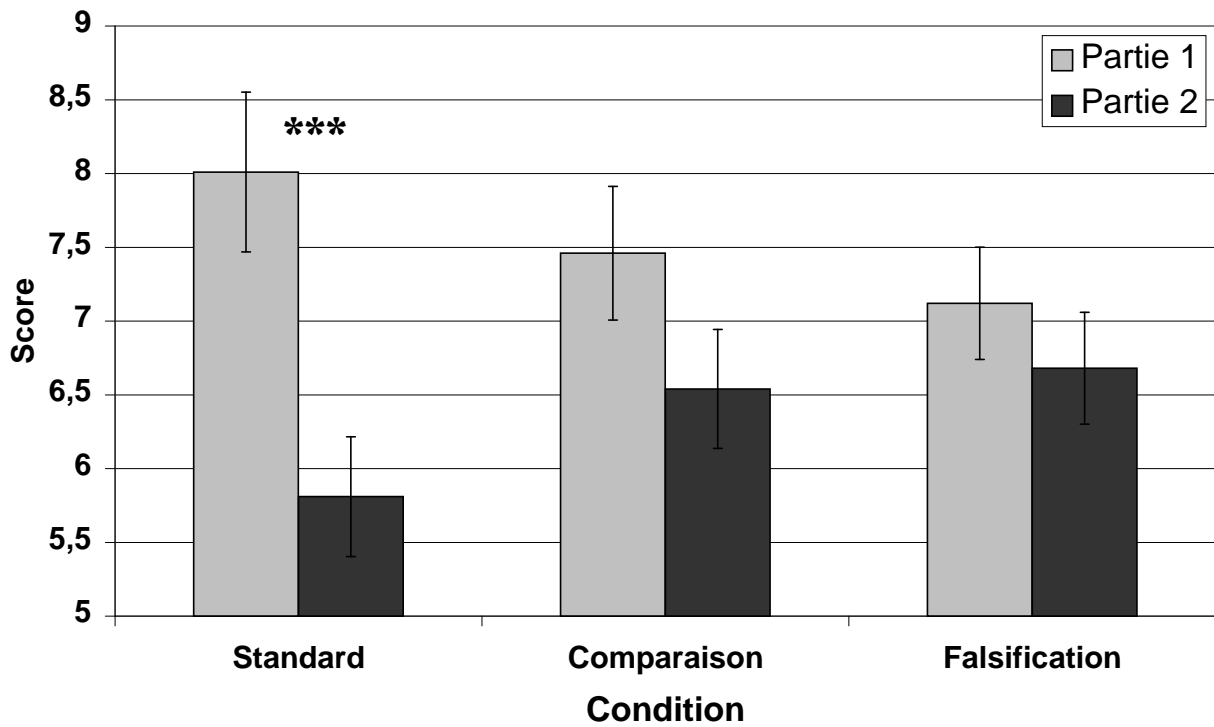
FIGURE 4.3 – Nombre de participantes (pourcentage) par niveau de performance au MRT.

Instructions et décours de la performance. Une ANOVA mixte *Instructions* \times *Section* répétée sur le second facteur montre un effet principal de la section du test, $F(1, 275) = 25.04$, $p < .001$; $\eta^2 = .083$, la seconde partie étant significativement moins bien réussie ($M = 6.32$, $SD = 3.89$) que la première ($M = 7.24$, $SD = 4.38$). Conformément à nos attentes, nous observons une interaction significative *Instructions* \times *Section*, $F(2, 274) = 4.07$, $p < .01$; $\eta^2 = .033$, sans effet principal des instructions, $F < 1$ (cf. fig.4.4 p. 136). Afin de préciser cette interaction, nous avons réalisé des contrastes orthogonaux dans chaque section du test, et ainsi comparé : (1) la performance en condition *standard* avec les performances dans les deux autres conditions, et (2) les performances dans les conditions *falsification* et *comparaison*. Ces contrastes se sont révélés non significatifs. En revanche, comme attendu, une série de test t selon les critères de Bonferroni³ (seuil à $\alpha = .016$) montre que l'on n'observe de différence significative, $t(72) = 5.18$, $p < .001$, entre les deux parties du test qu'en condition *standard* (cf. tableau 4.2). Il semble donc bien que la falsification du stéréotype, qu'elle soit explicite ou indirecte — *via* une comparaison avantageuse —, puisse prévenir la baisse de performance que l'on observe par ailleurs dans la condition *standard*.

TABLE 4.2 – Performance des étudiantes aux deux parties du MRT en fonction des instructions.

		Instructions		
		Standard	Comparaison	Falsification
<i>Score partie 1</i>	<i>M</i>	7.99	7.46	7.12
	<i>SD</i>	4.32	4.37	4.02
<i>Score partie 2</i>	<i>M</i>	5.75	6.54	6.68
	<i>SD</i>	3.47	3.89	4.01

3. Il n'existe pas de procédure particulière recommandée pour le calcul des contrastes dans les plans à mesure répété, sinon d'appliquer une correction de Bonferroni, qui consiste en la division du seuil de significativité choisi par le nombre de contrastes envisagés. Le défaut de cette procédure est d'être très conservatrice, et ce d'autant plus que le nombre de comparaison augmente.



*** : $p < .001$

FIGURE 4.4 – Interaction *Instructions* \times *Section* sur l'échantillon total.

Questionnaire

Les réponses aux questionnaires étaient analysés à l'aide d'une ANOVA simple avec le type d'instruction en facteur. Ce questionnaire était soumis aux participantes après le MRT, et leur réussite à ce test a pu influencer les réponses à ces questions. Aussi, le score obtenu au MRT est introduit en covarié dans l'analyse. Dans un soucis d'économie nous ne présentons les analyses que sur l'échantillon global, car les résultats ci-dessus ont montré que la distinction échantillon global / réduit n'est pas très pertinente pour la présente étude.

Stéréotype. Dans cette partie du questionnaire, les participantes répondaient à des questions du type : *Il est possible qu'en matière de compétences spatiales, les femmes aient d'une façon générale moins de facilités que les hommes* : (1) *Absolument pas d'accord* / ... / (6) *Absolument d'accord*. Les quatre questions portant sur l'adhésion au stéréotype

ont été agrégées (moyenne) en un indice unique (Cronbach $\alpha = .78$). Cependant, aucun effet significatif des instructions sur l'adhésion n'apparaît, $F(2, 273) = 1.56$, *n.s.*⁴. Les participantes se montrent incertaines à l'égard du stéréotype ($M = 3.55$, $SD = .77$, le milieu de l'échelle étant à 3.5 — one-sample t-test *n.s.*, $t(277) = .80$). Elles ne le rejettent pas clairement à l'inverse de ce que nous avons pu observer chez les collégiennes (cf. étude 1, p. 106).

Auto-évaluation. L'échelle d'auto-évaluation consistait en la question suivante : *Comment jugez-vous votre performance sur ce test (avec les figures) ? : (1) Médiocre /.../ (5) Excellente.* Les participantes s'auto-évaluent négativement, un test t pour échantillon unique indique que la moyenne des réponses à l'échelle d'auto-évaluation est significativement inférieure à 3, $t(277) = -14.67$, $p < .001$. ($M = 2.25$, $SD = .85$), il n'y a aucun effet des instructions sur l'auto-évaluation, $F(2, 274) = 2.32$, *n.s.* L'auto-évaluation comparative avec les autres étudiantes est également négative, le test- t pour un échantillon indique que le score global d'auto-évaluation comparative avec les autres étudiantes ($M = 2.50$, $SD = .69$) est significativement inférieur à 3, $t(277) = -11.92$, $p < .001$. Il n'y a aucun effet des instructions, $F(2, 274) = 1.31$, *n.s.* L'auto-évaluation comparative avec les hommes (étudiants) est également négative, le test- t pour un échantillon indique que ce score ($M = 2.38$, $SD = .76$) est significativement inférieur à 3, $t(277) = -13.68$, $p < .001$. Il n'y pas d'effet des instructions, $F(2, 274) = 2.51$, *n.s.* Enfin les participantes s'auto-évaluent positivement par rapport aux personnes âgées ($M = 3.71$, $SD = .83$), un test t pour échantillon unique indique que ce score est significativement supérieur à 3, $t(277) = 14.38$, $p < .001$. Il n'y pas d'effet des instructions, $F(2, 274) < 1$. Nous constatons que les auto-évaluations recueillies après le test vont dans le même sens que dans l'étude 2 où seul un questionnaire était proposé aux participants.

4. Pour cette analyse et toutes celles qui suivent, la performance des participantes au MRT a été contrôlée en l'introduisant comme covarié. Sauf lorsque mentionnée, cette procédure ne modifie en rien le profil ou la significativité des résultats rapportés.

Comparaison intersexe. Les participantes devaient répondre à la question : *Sur ce test, comment situeriez-vous la performance des participants de même sexe que vous par rapport à celle des participants de sexe opposé ? : (1) Très inférieure /.../ (5) Très supérieure.* Globalement les participantes estiment que les femmes obtiennent au MRT des scores inférieurs à ceux des hommes ($M = 2.84$, $SD = .52$), un test-t pour un échantillon indique que le score moyen sur cette échelle est significativement inférieur à 3, $t(277) = -5.16$, $p < .001$. Il n'y a pas d'effet des instructions $F(2, 274) = 1.34$, *n.s.*. Ainsi, si nous avons constaté précédemment que les participantes manifestaient de l'incertitude sur l'échelle d'adhésion au stéréotype, elles semblent ici accepter plus nettement l'idée d'une infériorité des femmes au MRT.

Comparaison inter-âge. Les participants devaient répondre à la question : *Sur ce test, comment situeriez-vous la performance des personnes de votre âge par rapport à celle des personnes âgées ? : (1) Très inférieure /.../ (5) Très supérieure.* Globalement les participantes estiment que les sujets jeunes obtiennent des scores supérieurs à ceux des sujets âgés au MRT ($M = 4.08$, $SD = .67$), un test-t pour un échantillon indique que le score moyen est significativement supérieur à 3, $t(277) = 11.87$, $p < .001$. Il n'y a pas d'effet des instructions $F(2, 274) < 1$. Les attentes des participantes vont donc dans le même sens que les attentes exprimées par les étudiantes dans l'étude 2.

Difficulté. Afin de mesurer la perception de la difficulté du MRT par les participants, ceux-ci répondaient à la question : *Que diriez-vous du niveau de difficulté de la première série d'exercices ? : (1) Très facile /.../ (5) Très difficile.* Globalement les participantes jugent le test plutôt difficile ($M = 3.34$, $SD = .82$), un test-t pour un échantillon indique que la moyenne de la comparaison inter-âge est significativement supérieure à 3, $t(277) = 6.88$, $p < .001$. L'analyse révèle un effet significatif des instructions sur la percep-

tion de la difficulté⁵, $F(2, 273) = 3.28$, $p < .05$; $\eta^2 = .023$, mais de manière surprenante, si le test est jugé plutôt difficile ($M = 3.15$, $SD = .89$) dans la condition *standard*, les participantes le jugent encore plus difficile dans les deux autres conditions (*falsification* : $M = 3.43$, $SD = .82$, *comparaison* : $M = 3.38$, $SD = .75$). Pourtant nous aurions pu attendre que les sujets placés dans les conditions les plus favorables à l'égard du stéréotype jugent le test plus facile qu'en condition standard.

4.3.3 Discussion

L'objectif de la présente étude était de reproduire chez des étudiantes le phénomène observé dans l'étude 1, c'est-à-dire l'effet positif de la falsification du stéréotype sur la réussite en RM. Cet objectif n'a été que partiellement atteint. Nous avons bien montré que lorsque les participantes recevaient des instructions *standards* leur performance baissait entre les deux parties du MRT conformément à nos précédentes observations quant à la dynamique du *stereotype threat*. Nous avons montré également que lorsque le stéréotype était falsifié explicitement (condition de *falsification*) ou indirectement (condition de *comparaison*), cette baisse n'était pas observée, suggérant que ces inductions étaient efficaces pour annuler les effets négatifs du stéréotype. La condition de comparaison, introduite dans cette étude 3, peut donc de ce point de vue être considérée comme équivalente à la falsification explicite de par ses effets sur la performance. Gardons cependant à l'esprit que dans cette étude 3, la performance globale (toute section du test confondues) ne diffère pas d'une condition à l'autre. Or la littérature conduirait à attendre une performance globale inférieure en condition standard relativement aux deux autres conditions, ce n'est manifestement pas le cas dans cette étude.

Une autre faiblesse tient à la taille de notre échantillon relativement à l'étude 1. Nous ne disposons *grosso modo* que de trois fois moins de sujets que dans l'étude 1, ce qui

5. L'effet n'est plus significatif mais juste tendanciel, $F(2, 275) = 2.77$, $p = .064$; $\eta^2 = .020$, lorsque le score obtenu au MRT n'est plus introduit comme covarié dans le modèle statistique. Il est évident que la difficulté perçue est la dimension sur laquelle le score des participantes a le plus d'effet.

limitait beaucoup les possibilités de découpage de l'échantillon pour le calcul des distributions de performance. De manière générale, le MRT étant très difficile, peu de sujets atteignent les hauts scores, d'où la nécessité d'un échantillon de base très important pour disposer finalement d'un sous-échantillon suffisamment grand au niveau des hauts scores. Ici, moins de 35 sujets étaient répartis dans les niveaux *bon* et *très bon*, toutes conditions confondues, ce qui est peu. À ces niveaux de performance, les comparaisons entre conditions portaient donc sur des effectifs très réduits, ce qui conduit donc à douter de leur fiabilité. Un autre problème assez proche concerne la définition même des niveaux de performance. Dans l'étude 1, la présence de garçons dans l'échantillon permettait de définir en quelque sorte un « niveau de référence », qui permettait de situer de manière lisible les performances des filles. Ici, du fait de l'absence d'hommes, nous avons dû situer les performances des femmes par rapport à un découpage « absolu » de l'échelle des scores au test, qui n'étaient visiblement pas adapté à la forme des distributions de performances, le nombre de femmes atteignant les hauts score étant très faible.

Par ailleurs, nous pouvons envisager plusieurs arguments en faveur de l'idée que les effets du stéréotype n'ont pas été particulièrement importants dans cette étude. Tout d'abord, il existe une différence de taille entre la présente étude et celle menée chez les collégiens : les conditions de mixité lors de la passation. Dans chaque classe de collège, le nombre de garçons et de filles était sensiblement équilibré. Or, concernant la passation des étudiantes, le nombre d'hommes présents au moment de recevoir le test était nettement marginal, sinon nul. Or il a été montré que la proportion d'individus du sexe opposé était un des éléments participant à l'activation du stéréotype au cours d'un test (Sekaquaptewa & Thompson, 2003). On peut donc envisager le fait que les étudiantes aient été très majoritaires lors de la passation (environnement non mixte) a pu réduire l'impact de l'interaction entre leur identité de genre et l'activation du stéréotype, et ainsi atténuer voir empêcher la survenue d'une interférence avec le stéréotype.

D'autre part, dans le même sens, il est possible que des étudiantes en psychologie aient été moins sensibles à l'interférence du stéréotype de par leur moindre identification au domaine. Les étudiantes en psychologie n'ont *a priori* aucune réputation de compétence à défendre dans les domaines associés à la RM (sciences, géométrie, etc.), car ces compétences ne leur sont pas typiquement attribuées. Or l'effet du stéréotype est particulièrement critique chez les individus fortement identifiés au domaine évalué, pour lesquels l'évaluation constitue justement un enjeu pour leur réputation et leur identité. Un tel enjeu a peut-être ici fait défaut. Aussi, si les étudiantes en psychologie constituaient une population facilement mobilisable en nombre, elle n'appartenaient peut-être pas au groupe le plus sensible aux effets que nous avons cherché à mettre à jour.

4.4 Conclusion

L'objectif des deux études présentées dans ce chapitre était de s'assurer d'une part de l'existence d'un stéréotype de genre dans la perception de la RM, et d'autre part que ce stéréotype est de nature à entraver les capacités des femmes en RM. L'étude 2 a clairement mis en évidence l'existence d'un stéréotype chez les hommes, alors que les femmes semblent lui résister davantage, au moins dans le cas où la question est abordée directement. En effet des questions plus indirectes montrent que les étudiantes nourrissent des attentes compatibles avec le stéréotype. Concernant la mise en évidence du rôle du stéréotype dans le MRT, nous avons montré qu'une intervention qui ne faisait pas directement référence au stéréotype (comparaison avec des personnes âgées) peut avoir un effet positif similaire à la falsification explicite du stéréotype. Ces deux interventions conduisent à un maintien des performances entre les deux sections du test, tandis que dans la condition classique de passation du MRT (condition standard), nous constatons une baisse du score dans la deuxième section.

Néanmoins l'effet positif de nos interventions est trop faible pour aboutir à une augmentation notable du score des femmes. Ce résultat en demi-teinte peut s'expliquer par certains éléments méthodologiques ou liés à la situation de test. La population testée, de part sa faible identification au domaine scientifique et visuo-spatial (connus pour être liés) n'était peut être pas la plus sensible à l'intervention du stéréotype.

Dans le prochain chapitre nous présentons une étude expérimentale en laboratoire auprès d'étudiantes inscrites dans des cursus littéraire ou scientifique, afin de comparer deux populations qui n'ont pas *a priori* la même identification au domaine visuo-spatial. Mais surtout, cette étude conduite en laboratoire offre la possibilité d'intégrer une dimension chronométrique (analyse des temps de réponse) pour mieux apprécier la nature des effets liés à l'intervention d'un stéréotype.

Chapitre 5

Impact du stéréotype sur les mécanismes énergétiques

5.1 Objectifs de l'étude 4

Notre objectif dans ce chapitre est non seulement de valider le paradigme de falsification sur une version informatisée d'une tâche de RM en trois dimensions, mais aussi de montrer combien le recueil et l'analyse des temps de réponse peut éclairer les mécanismes sous-jacents à l'interférence du stéréotype. Nous rappelons que l'effet du stéréotype est surtout étudié via des tâches de haut niveau (tests verbaux, de raisonnement, de mathématique, etc.) généralement administré dans un format « papier-crayon », qui ne permettent pas une décomposition précise des processus cognitifs mis en jeu, à la différence par exemple des techniques employées en chronométrie mentale. En fait, les mécanismes sous-jacent au *stereotype threat* sont d'ordinaire appréhendés par la mesure de facteurs (anxiété, charge mentale, capacité en mémoire de travail, etc.) dont on teste l'implication grâce à des analyses de médiation. La synthèse des résultats qui ont permis d'isoler des médiateurs de l'intervention du stéréotype a permis à Schmader et coll. (2008) de propo-

ser un modèle intégré de ce phénomène. Nous pensons qu’une approche chronométrique (analyse des temps de réponse) est possible, et permettrait de cerner quelles composantes sont impliquées dans la baisse de performance observée. Pour soutenir cette approche chronométrique, nous nous appuyerons sur les prédictions d’un modèle général du stress et de la performance (Sanders, 1983) qui indique comment les conditions expérimentales pourraient affecter les temps de réponse.

5.2 Modèle intégré du *stereotype threat*

5.2.1 Médiateurs du *stereotype threat*

Afin d’introduire les modèles explicatifs du *stereotype threat*, nous allons dans un premier temps passer rapidement en revue les études qui ont cherché à isoler les médiateurs responsables de la baisse de performance due à une intervention du stéréotype. Parmi les différents médiateurs potentiels, l’anxiété est le facteur qui a peut-être été le plus testé. En effet, la formulation même du phénomène par Steele et Aronson, « *menace* du stéréotype », fait immédiatement référence à l’anxiété, que nous pouvons définir comme la réponse d’un individu à une sollicitation de l’environnement perçue comme une menace. Pour Delignières (1993), dans un contexte évaluatif, l’anxiété est une réponse émotionnelle consécutive à la perception par le sujet d’un écart entre la difficulté d’une tâche et sa capacité à réussir. Mais cet auteur rappelle que cette perception n’est pas suffisante pour rendre la situation anxiogène, il faut également qu’il soit important pour le sujet de réussir. Ainsi, l’anxiété dans une situation d’évaluation repose sur deux éléments : l’*incertitude* quant au résultat et l’*enjeu* que représente ce résultat. Delignières conclut que l’incertitude relève de la « confrontation tâche-sujet », tandis que l’enjeu « dépend plutôt d’une interaction contexte-sujet » (Delignières, 1993, p. 236). Il est clair que le *stereotype threat* fait intervenir ces deux composantes : la confrontation à la difficulté du test,

souvent augmentée d'une pression temporelle, crée l'incertitude sur l'issue du test, tandis que le risque d'échouer et donc de confirmer le stéréotype, crée l'enjeu. De ce point de vue, l'anxiété semble donc bien le médiateur le plus probable de la baisse de performance provoquée par l'intervention du stéréotype.

Cependant, les différentes études qui ont testé la médiation de l'effet du stéréotype par l'anxiété divergent dans leurs conclusions. Les résultats de Steele et Aronson (1995) vont dans le sens d'un rôle médiateur de l'anxiété, de même que Spencer et coll. (1999), quoique de l'aveu de ces derniers, la relation de médiation soit faible. Osborne (2001) a pu établir une corrélation entre performance et anxiété, mais n'a pu garantir qu'un stéréotype était bien intervenu lors de la passation. Stone et coll. (1999) ont conclu que la corrélation qu'ils observaient entre l'anxiété et la performance n'atteignait pas un niveau suffisant pour permettre le calcul d'une médiation. Aronson et coll. (1999) quant à eux n'ont pu déceler d'effet de l'anxiété, ce qui peut s'expliquer selon eux par le manque de fiabilité des mesures de l'anxiété (échelles auto-rapportées), et d'autre part parce que le nombre de médiateurs potentiels de l'effet du stéréotype est important, et qu'il est possible que tous jouent un rôle à des niveaux divers, voire interagissent entre eux.

Concernant les autres médiateurs possibles, Blascovich, Spencer, Quinn, et Steele (2001) ont montré une élévation de la pression sanguine chez des Afro-américains au cours d'une situation de *stereotype threat*, ce qui a été interprété comme la preuve matérielle d'une réaction anxieuse. Mais outre le fait que ces auteurs n'ont procédé à aucune analyse de médiation, l'on sait par ailleurs que l'anxiété « cognitive » — supposée à l'origine de l'interférence du stéréotype — est indépendante de l'anxiété « somatique », dont la manifestation se traduit effectivement par une modification de certaines variables physiologiques, mais qui ne semble pas avoir de conséquences négatives sur la performance (Delignières, 1993). D'autre part, lorsque Croizet et coll. (2004) ont mesuré une autre variable physiologique, la variabilité cardiaque, ils ont interprété cette mesure comme un indice de la charge mentale, et ont ainsi montré que cette charge médiatisait l'effet du

stéréotype. Enfin, nous pouvons évoquer des résultats montrant l'implication de facteurs tels que le niveau d'activation du stéréotype lui-même (Davies et coll., 2002), le niveau d'éveil (*arousal*, Ben-Zeev, Fein, & Inzlicht, 2005 ; O'Brien & Crandall, 2003), le nombre de pensées négatives (Cadinu et coll., 2005), et la capacité en mémoire de travail (Beilock et coll., 2007 ; Schmader & Johns, 2003). L'anxiété ne peut donc jouer un rôle exclusif pour expliquer l'interférence du stéréotype, et modéliser ce phénomène nécessite donc de spécifier les interactions entre les différents facteurs impliqués. Schmader et coll. (2008) ont proposé un tel modèle.

5.2.2 Le modèle intégré de Schmader et coll. (2008)

La figure 5.1 présente un schéma qui décrit les interactions des différents mécanismes qui, selon Schmader et coll., conduisent à une baisse de la performance suite à l'activation du stéréotype. Pour un sujet fortement identifié au domaine évalué, l'activation du stéréotype alliée à la difficulté du test fait de ce dernier une situation anxiogène. Ce stress provoque d'une part une réponse physiologique (augmentation du niveau d'éveil), et d'autre part l'activation d'un processus de contrôle sensible aux éléments du test interprétés comme des indices permettant de réduire l'incertitude du sujet quant à l'issue du test (réussite ou échec). Ce processus, biaisé par l'activation du stéréotype, va être responsable d'une augmentation du nombre de pensées négatives parasites associées au test. Pour inhiber ces pensées négatives, le sujet devra fournir un effort particulier, pour mettre en œuvre un processus « d'antiparasitage ». L'augmentation du niveau d'éveil, l'activation du processus de contrôle et celle du processus d'antiparasitage consomment (par des voies différentes) une partie des ressources de la mémoire de travail, ce qui aboutit à une baisse de performance sur des tâches qui nécessitent l'intervention d'un contrôle cognitif. Comme le stéréotype semble pouvoir interférer également avec des tâches où la mémoire de travail n'intervient pas (processus automatiques, par exemple des tâches sportives faisant appel à la mémoire procédurale, cf. Beilock, Jellison, Rydell, McConnell,

& Carr, 2006), Schmader et coll. envisagent également l'existence d'un lien direct entre le processus sensible aux caractéristiques de la tâche (indices) et les processus automatiques. Mais cela ne concerne pas notre problématique, la RM reposant directement sur la manipulation de l'information en mémoire de travail.

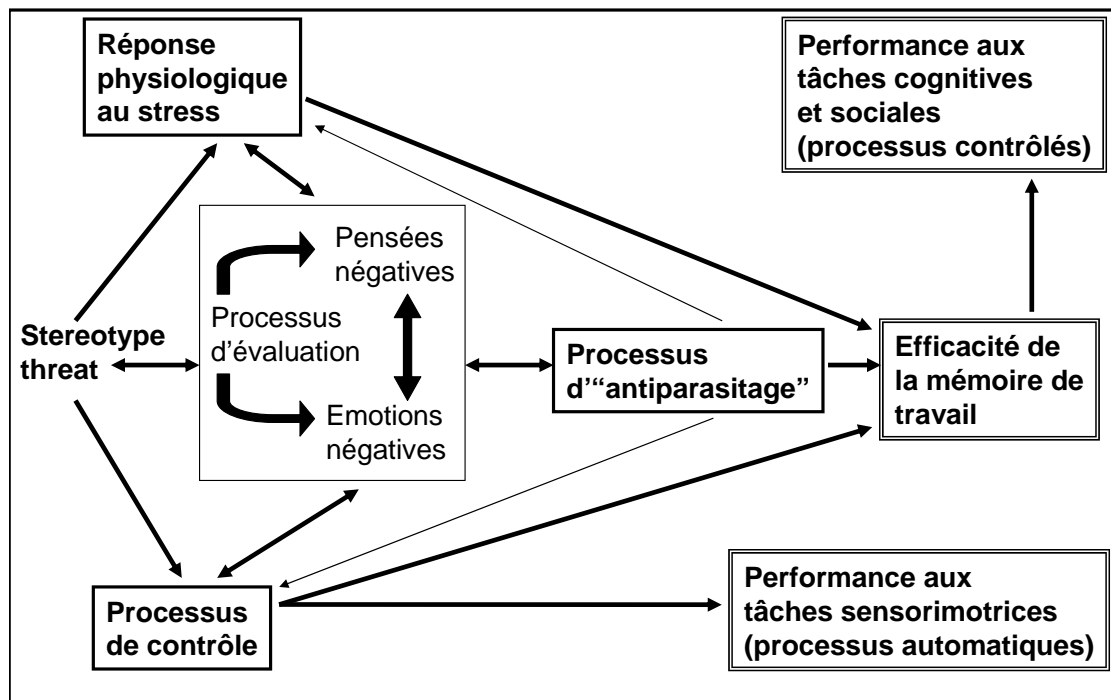


FIGURE 5.1 – Modèle intégré de Schmader et coll. (2008) montrant les différents mécanismes impliqués dans la baisse de performance due au *stereotype threat*.

Le modèle de Schmader et coll. permet de mettre en relation les différents facteurs impliqués dans le *stereotype threat* de manière cohérente. Cependant, il ne permet pas de prédiction très précise concernant l'impact du phénomène sur les temps de réponse (TR). Tout au plus, s'il prédit une baisse des ressources disponibles en mémoire de travail, nous pouvons en déduire un ralentissement de la vitesse de traitement de l'information spatiale, notamment au niveau des angles de rotation les plus importants, qui mobilisent une part plus importante des ressources. Or Sanders (1983) a proposé un modèle du stress et de la performance qui tout en étant compatible avec le modèle de Schmader et coll., permet de faire des hypothèses beaucoup plus précises sur l'évolution des TR en fonction des situations de stress.

5.3 Le modèle du stress de Sanders (1983)

L'objectif de Sanders (1983) était de réunir deux approches permettant d'expliquer la performance du traitement de l'information. Une première approche, qualifiée de « computationnelle », décrit ce traitement de manière séquentielle, en une succession d'étapes. L'autre approche, « énergétique », considère les ressources cognitives comme une quantité limitée plus ou moins disponible en fonction de différents facteurs liés à la situation, l'état du sujet, etc. Sanders (1983) a proposé un modèle de couplage de ces deux types de processus, les mécanismes énergétiques modulant l'efficacité du traitement computationnel de l'information, le tout sous la supervision d'un processus d'évaluation de la performance en fonction des feedbacks reçus et des buts (motivation) du sujet.

Le modèle de Sanders se décompose en trois niveaux (cf. fig. 5.2). Le premier correspond aux différentes étapes de traitement séquentiel de l'information, et Sanders a simplement choisi d'intégrer la chaîne de traitement proposée par Sternberg (1969) : *pré-traitement*, *extraction des caractéristiques*, *sélection de la réponse* et *ajustement moteur*. Ce niveau permet d'expliquer la performance par l'efficacité du traitement opéré à chaque étape de la chaîne, qui peuvent être altérées spécifiquement par certains facteurs (intensité du stimulus, qualité du signal, etc.). Le second niveau du modèle regroupe les processus énergétiques, et a été inspiré par la théorie de l'*effort* de Kahneman (1973), ainsi que par plusieurs modèles neuro-anatomiques (Gray, 1971 ; Pribram & McGuinness, 1975 ; Routenberg, 1968), ou encore par les trois composantes de l'attention proposées par Posner et Boies (1971) — à savoir vigilance, attention sélective et traitement conscient. Pour sa part, Sanders distingue trois composantes qui sont : *niveau d'éveil (arousal)*, *activation* et *effort*. Chacune de ces composantes énergétiques est couplée avec une des étapes du traitement de l'information (cf. fig. 5.2).

Le point central du modèle de Sanders est la manière dont les processus énergétiques interagissent, afin de rendre compte de l'évolution des performances en fonction du stress.

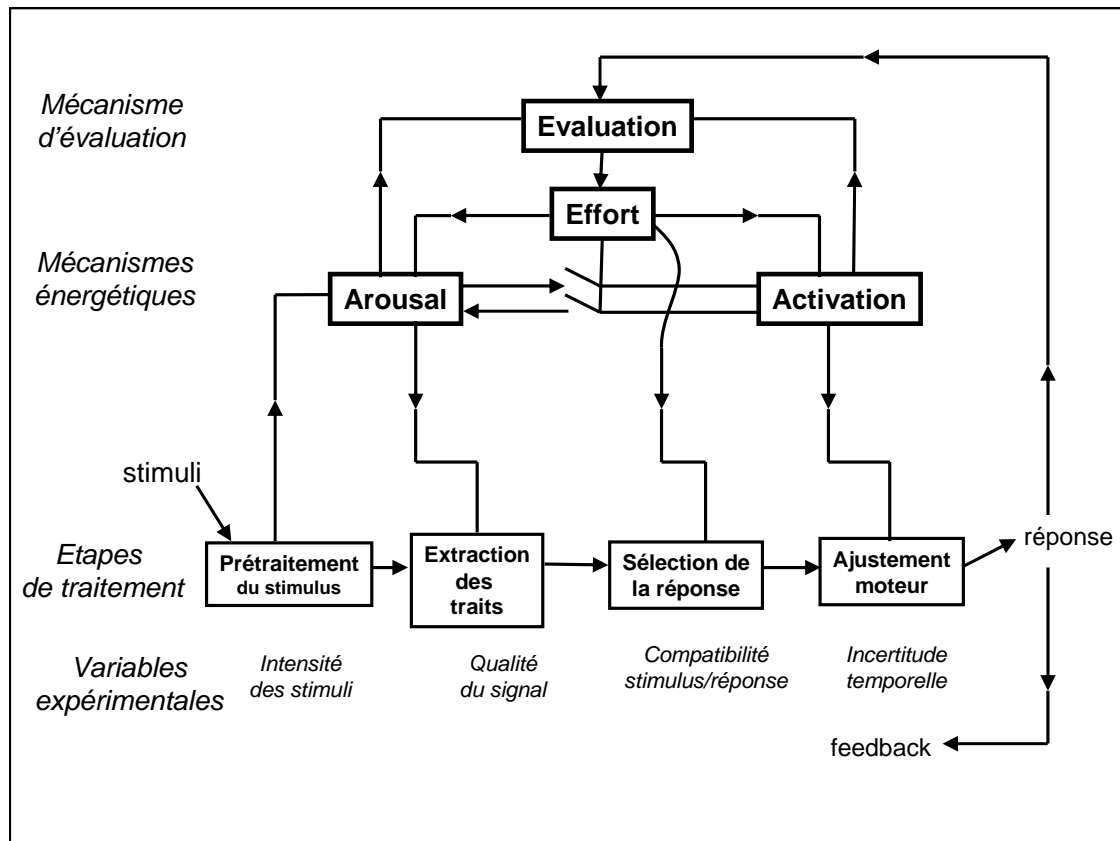


FIGURE 5.2 – Modèle du stress et de la performance de Sanders (1983) montrant l'interaction entre les différentes étapes du traitement cognitif de l'information, et les processus énergétiques modulant l'efficacité de ce traitement sous la supervision d'un processus d'évaluation.

Ces processus ont pour fonction de compenser le déséquilibre engendré par des perturbations de différentes natures et qui peuvent avoir un impact sur les performances, en gérant l'allocation des ressources énergétiques. Les perturbations peuvent être d'origine environnementale (p. ex. température ambiante, etc.), propres à la tâche (dégradation des stimuli, etc.) ou encore propre au sujet lui-même (manque de sommeil, etc.). Alors que le niveau d'éveil (*arousal*) et l'activation sont des mécanismes énergétiques assez basiques, le premier intervenant au niveau de la perception des stimuli et le second au niveau de la réponse motrice à ces stimuli, l'effort est lui un processus contrôlé, couplé au stade de sélection de la réponse aux stimuli. L'effort a pour rôle de coordonner les deux autres mécanismes énergétiques pour éviter des déséquilibres. C'est l'incapacité de l'effort à rééquilibrer / compenser le déséquilibre en cas de perturbation trop importante (par exemple tâche trop difficile, où conditions d'exécution trop dures) qui selon Sanders

définit le stress. Cela ne signifie pas qu'un effort important génère du stress en lui-même : en fait c'est la présence d'un risque de rupture de l'équilibre entre les mécanismes énergétiques (ou simplement la survenue de cette rupture) qui définit le stress. Pour preuve, il est bien connu que le stress suit une courbe à optimum (« U » inversé, cf. Yerkes & Dodson, 1908) : une sollicitation trop faible (par ex. tâche facile, mais monotone) peut créer un stress de la même manière qu'une sollicitation trop importante.

Le dernier niveau du modèle enfin consiste en un système de régulation (contrôle cognitif) de l'effort. Il est constitué par un mécanisme d'évaluation qui reçoit ses entrées des deux mécanismes énergétiques de base (éveil cortical et activation) ainsi que de la réponse fournie par le sujet et du feedback reçu. Il détermine ainsi dans quelle direction l'effort doit être dirigé ou encore si les mécanismes énergétiques de base doivent être dissociés, afin d'éviter que la trop forte perturbation de l'un se propage à tout le système. Sanders passe en revue un grand nombre de situations qui peuvent provoquer de telles perturbations (tâches monotones, privation de sommeil, absorption d'excitants, etc.), mais pour notre propos, nous ne nous focaliserons que sur les effets de l'interférence du stéréotype avec les mécanismes énergétiques.

5.3.1 Modèle énergétique et interférence du stéréotype

L'anxiété provoquée par l'intervention du stéréotype peut être vue comme un facteur de déséquilibre des mécanismes énergétiques contre lequel le sujet devra fournir un effort. En effet, une augmentation de l'activation et de l'*arousal* peuvent correspondre à une augmentation de l'anxiété respectivement somatique et cognitive¹. Ben-Zeev, Fein, et Inzlicht (2005) ont présenté des résultats permettant d'attribuer les effets négatifs de l'intervention du stéréotype à une augmentation du niveau d'éveil au delà de l'optimum. Après avoir activé un stéréotype de sexe en mathématiques, ces auteurs ont proposé deux

1. Cette distinction des deux composantes de l'anxiété n'est de plus pas très critique dans notre cas, certains auteurs ne distinguent que deux composantes, éveil / activation (sans distinction) d'un côté, et effort de l'autre (cf. Humphreys & Revelle, 1984).

types de tests (facile vs. difficile) sans rapport avec les mathématiques à des femmes, et ont constaté une baisse de performance au test difficile alors que les performances étaient supérieures au test facile, en comparaison avec un groupe de femmes où le stéréotype n'était pas activé. C'est un profil de résultat compatible avec l'idée d'une augmentation de l'*arousal*, car ce dernier a typiquement un effet positif pour les tâches faciles et négatif au delà d'un certain niveau de difficulté. Ces données invalident les explications reposant sur l'hypothèse d'un abaissement général des capacités cognitives provoqué par le stéréotype.

Le modèle de Sanders est non seulement compatible avec les résultats expérimentaux concernant le *stereotype threat*, mais également avec le modèle proposé par Schmader et coll. (2008). Pour ces derniers, l'activation du stéréotype crée une situation anxiogène qui induit une réponse physiologique et une sensibilité accrue aux indices « menaçants » dans la situation de test. Pour modérer ces deux mécanismes, un processus d'évaluation déclenche un processus de régulation (« anti-parasitage ») afin de maintenir la performance. On peut établir une correspondance avec les éléments du modèle de Sanders, qui postule également l'existence d'un processus d'évaluation contrôlant un processus de régulation (ici appelé effort). Ce dernier agit sur des mécanismes énergétiques associés à une réponse physiologique face à une situation anxiogène (*arousal* et activation) et couplés avec des processus computationnels d'extraction et de traitement de l'information. Dans les deux modèles, l'ensemble du mécanisme de régulation est soumis à un contrôle, la principale différence étant que si Schmader et coll. (2008) mettent l'accent sur l'implication de la mémoire de travail, Sanders détaille plutôt les différentes étapes de traitement de l'information. Cependant on peut très bien envisager une implication de la mémoire de travail dans les processus de contrôle de l'effort, d'évaluation, et de sélection de la réponse considérés par le modèle de Sanders.

Enfin, nous considérons le modèle de Sanders très pertinent vis-à-vis du type de problèmes de RM que nous allons présenter aux sujets dans cette étude. Nous allons recueillir des TR de choix, les sujets devront sélectionner une réponse correcte face à deux propo-

sitions qui leur seront présentées (deux objets 3D à comparer). Or le modèle de Sanders intègre un modèle de sélection de la réponse en fonction du stimulus présenté, il est donc tout indiqué pour appréhender les effets du stéréotype dans une telle tâche. Sanders livre par ailleurs quelques indications sur l'impact respectif de chaque composante du modèle — computationnelle ou énergétique — sur les TR. Le temps de réponse à un problème particulier (essai) serait composé du temps nécessaire au traitement des stimuli (composante computationnelle) — chaque étape correspondant à une durée de traitement précise —, auquel s'ajoute une durée dépendante de la composante énergétique. Selon Sanders, un facteur qui impacterait la composante computationnelle (par exemple un changement de stratégie) modifierait de manière *systématique* une (ou plusieurs) étape du traitement des stimuli. Ceci aurait pour conséquence une modification dans des proportions identiques de la vitesse des réponses à une catégorie de stimuli donnée, ou autrement dit, une translation de l'ensemble de la distribution des TR. Par exemple, si le sujet opte pour une stratégie moins efficace, on devrait constater une augmentation comparable de tous les TR. Par contre un facteur qui impacterait la composante énergétique aurait des effets moins systématiques, car l'influence des processus énergétiques peut varier significativement d'un essai à l'autre, (cf. par exemple une fluctuation de l'attention au court du test). Plus les TR sont longs, et plus la part due aux processus énergétiques est importante — celle due à la composante computationnelle étant fixée et comparable d'un essai à l'autre. Aussi, c'est dans la partie de la distribution correspondant aux TR les plus longs que les effets énergétiques sont les plus visibles. Ceci nous permet d'avancer cette hypothèse : si l'intervention du stéréotype affecte non pas le traitement computationnel de l'information en lui-même mais les facteurs énergétiques (gestion et allocation des ressources cognitives), alors nous devrions observer un allongement des TR les plus longs dans la condition où le stéréotype est activé par rapport à une condition où il est falsifié. Par contre, nous ne devrions observer aucune différence entre les deux conditions pour les TR les plus courts.

Notre hypothèse portant sur un effet localisé au niveau des TR les plus longs, la méthode classique d'analyse des moyennes des TR n'est donc pas adaptée. Nous présenterons

et mettrons en œuvre une technique alternative : l'analyse *distributionnelle* des TR, qui consiste à modéliser la distribution des TR et à calculer les paramètres correspondants. Cette méthode d'analyse, qui suscite un intérêt croissant, permet par exemple de déceler, outre l'effet d'interférence, un effet de facilitation sur les items congruents d'une tâche de Stroop lorsque les méthodes d'analyses classiques (analyse de la tendance centrale des distributions de TR) ne sont pas assez discriminantes pour déceler un tel effet (Heathcote, Popiel, & Mewhort, 1991). D'autres auteurs ont pu grâce à ce type d'analyse mieux caractériser les profils des TR recueillis auprès d'enfants souffrants de déficits de l'attention et donc les distinguer des enfants sans problèmes attentionnels (Leth-Steensen, King Elbaz, & Douglas, 2000). Enfin, plus récemment, Davranche, Audiffren, et Denjean (2006) ont été les premiers à notre connaissance à procéder à une analyse distributionnelle en relation avec le modèle de Sanders, mais dans une optique très différente de la nôtre, sans chercher à inférer la nature énergétique ou computationnelle des processus impliqués à partir de la forme des distributions des TR recueillis. Ces auteurs s'intéressaient seulement à la mesure de l'effet de l'entraînement sur la forme des distributions des TR.

Pour tester notre hypothèse, nous avons conçu une version informatisée d'un test de RM. Afin de permettre le recueil de TR exploitables dans le cadre théorique envisagé, cette version informatisée ne pouvait inclure des problèmes de même type que ceux présentés dans le MRT de Vandenberg et Kuse (1978), où à chaque essai le sujet doit sélectionner deux réponses parmi quatre propositions. Nous avons donc créé des problèmes semblables à ceux proposés originellement par Shepard et Metzler (1971), où les objets sont présentés par paires. Les sujets doivent ainsi comparer deux objets 3D vus sous deux angles différents, et déterminer s'ils sont identiques ou différents (paradigme de *temps de réponse de choix*).

La présentation par paire n'est pas sans conséquence concernant les différences de sexe : il est connu que ce type de test montre moins systématiquement cette différence au niveau du taux de réponses correctes, ou alors avec une taille d'effet réduite. Rappelons que

Shepard et Metzler, qui ont créé le paradigme originel, n’avaient pas décelé de différence de sexe sur leur tâche — différence qu’ils ne cherchaient d’ailleurs pas spécialement à mettre en évidence. Ce résultat contraste avec la taille de la différence observée sur le MRT de Vandenberg et Kuse (1978), qui se sont pourtant inspirés du matériel de Shepard et Metzler. Selon Peters et Battista (2008), cela est probablement dû au fait que, dans le MRT, le nombre d’objets présentés simultanément est plus important, au point de peser plus encore sur les ressources cognitives alors disponibles. De plus, la pression temporelle est aussi plus forte dans le MRT — le temps donné est très court pour l’ensemble du test —, alors que dans le cas de présentation par paire, on attend simplement la réponse des sujets sans autre contrainte. Cela ne signifie pas nécessairement que le stéréotype ne joue aucun rôle au cours d’un test où les présentations ont lieu par paire : les capacités en RM des femmes peuvent très bien être entravées, ce qui les conduirait à répondre plus lentement (effet sur les TR), mais comme le temps n’est pas limité, cela n’a pas de conséquence sur le taux de réponses correctes. L’effet serait donc localisé sur les TR, et lorsque le stéréotype est falsifié, nous devrions alors observer chez les étudiantes les plus sensibles au phénomène un rapport vitesse-précision qui va dans le sens d’une plus grande efficacité, soit des réponses plus rapides pour une justesse au moins identique à celle observée dans les autres conditions.

Enfin, dans le chapitre précédent, alors que l’étude 2 montre bien l’existence d’attentes négatives concernant les résultats des femmes au test de RM, l’étude 3 ne montre pas d’effet positif de la falsification sur la performance totale. Un des points essentiels de la théorie de la menace du stéréotype est que la situation de test constitue un enjeu, ce qui est le cas pour les individus performants dans le domaine évalué. Les capacités en RM étant fortement corrélées avec la réussite dans les domaines mathématiques et scientifiques (Casey, Nuttall, et coll., 1995), nous pouvons envisager que l’effet bénéfique de la falsification du stéréotype sera plus visible chez des étudiantes de sciences plutôt que des étudiantes de lettres, moins performantes et de fait moins identifiées au domaine visuo-spatial *a priori*.

5.4 Méthode

5.4.1 Participants

Les participantes étaient recrutées par voie d’annonce sur des panneaux d’affichage, les invitant à participer à une expérience intitulée « Perception et contexte », avec une proposition d’indemnisation de cinq euros. Notre échantillon était composé de 77 étudiantes (âge moyen = 22.63) avec un bac littéraire ($N = 33$) ou scientifique ($N = 44$), aléatoirement affectées à la condition *standard* ou de *falsification*. Nous avons initialement testé 88 étudiantes, mais deux étudiantes ont quitté le test trop tôt suite à une mauvaise compréhension des instructions, et pour deux autres un problème informatique a interrompu la passation. Enfin un pré-traitement des résultats a révélé que 2 participantes pouvaient être suspectées d’avoir répondu au hasard (taux de réponses correctes proche de 50% pour tous les angles, et TR inférieurs à 2000 ms pour tous les angles également). Enfin, à l’issue de l’analyse distributionnelle, 5 étudiantes ont été retirées de l’échantillon car les tests d’ajustement (χ^2) indiquaient que les paramètres estimés ne rendaient pas compte de manière satisfaisante des distributions des TR recueillis auprès de ces sujets.

5.4.2 Procédure

Les participantes étaient accueillies individuellement par l’expérimentateur (de sexe masculin) et invitées à prendre place dans un box face à un écran et un clavier dont deux touches (« q » et « m ») avaient été couvertes d’un autocollant de couleur respectivement rouge et vert. Les sujets étaient alors invitées à lire les instructions qui s’affichaient à l’écran à leur rythme (défilement contrôlé par la touche « espace » du clavier), en présence de l’expérimentateur, qui répondait aux éventuelles questions. Deux exemples simples de problèmes étaient ensuite proposés, de manière à ce que l’expérimentateur puisse constater que les instructions étaient bien comprises. L’expérimentateur quittait alors le box, et le

sujet lançait un bloc de familiarisation de 60 essais, constitués de problèmes identiques à ceux proposés ultérieurement dans la tâche.

Chaque essai se déroulait comme suit : un écran vide était affiché pendant 1s, puis une croix de fixation apparaissait au centre de l'écran pendant 500 ms afin de capter l'attention du sujet, et enfin un problème était présenté jusqu'à ce que le sujet donne une réponse (cf. fig. 5.3 p. 157). Les problèmes consistaient en la présentation simultanée d'objets en 3D similaires à ceux créés par Shepard et Metzler (1971). Les sujets devaient déterminer si les deux objets étaient différents — auquel cas ils devaient appuyer sur la touche rouge — ou identiques (touche verte). Les paires différentes étaient constituées par des objets qui étaient l'image en miroir l'un de l'autre. Enfin il y avait trois angles de rotation possibles : zéro degré (pas de rotation), 60 degrés et 150 degrés. Nous avons évité des angles de 90 et 180 degrés car il y avait un risque, avec ces configurations, que des stratégies autres que la RM soient efficaces pour sélectionner la réponse, comme le « *flipping* » : la représentation mentale directe de l'image en miroir de l'objet après un basculement à 180 degrés (Kanamori & Yagi, 2002 ; Murray, 1997). Nous n'avons également proposé que trois angles différents afin de pouvoir pour chacun effectuer un nombre suffisant de mesures, comme le réclame l'analyse distributionnelle des TR. Ce choix nous permettait aussi de ne pas faire durer la tâche trop longtemps.

Après la phase de familiarisation, une première manipulation visait à activer l'identité de sexe des sujets. L'écran affichait l'image présentée figure 5.4 (p. 158), en même temps que les participantes recevaient l'instruction (écrite à l'écran) de cliquer sur le symbole correspondant à leur sexe. Même si seules des étudiantes étaient testées, le fait de demander de choisir entre les deux sexes suggérait la possibilité que des participants de sexe masculin étaient également testés, favorisant l'idée d'une éventuelle comparaison intersexe. La procédure choisie ici pour activer l'identité de sexe est ainsi beaucoup plus flagrante que celle que nous avons retenu pour les études précédentes, où les participants renseignaient simplement leur sexe au tout début du test. Ensuite, la phrase suivante s'affichait pendant

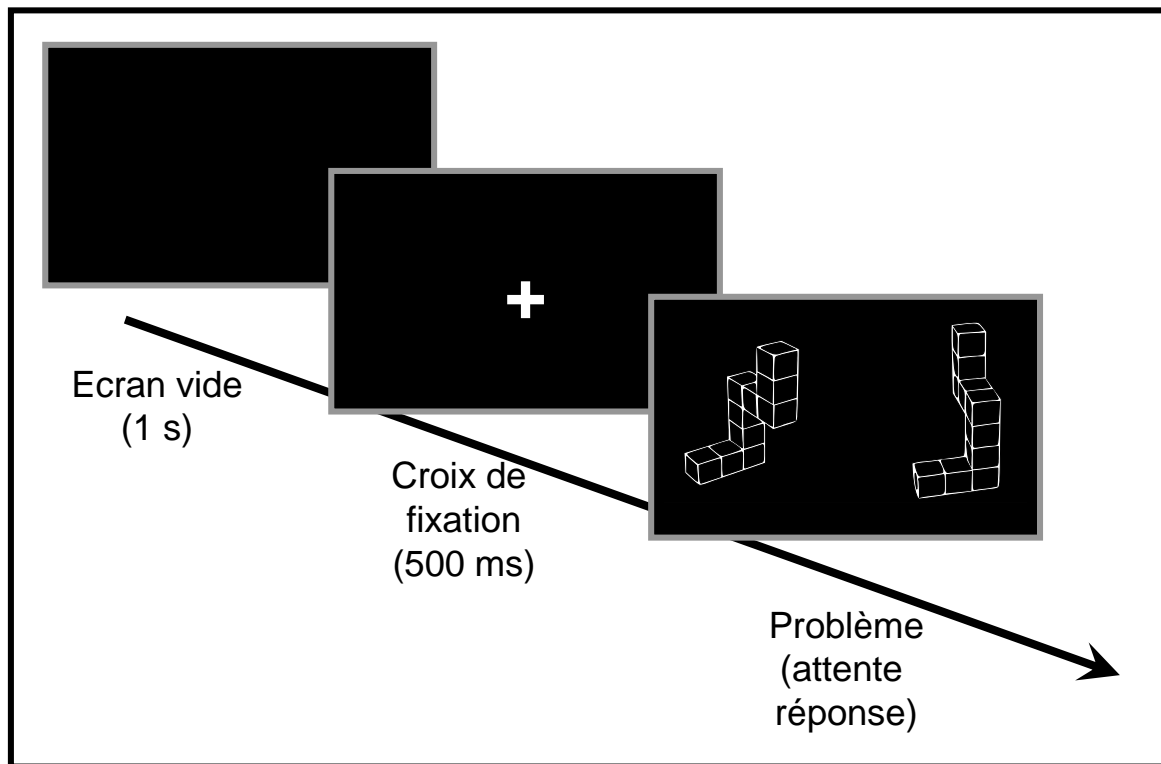


FIGURE 5.3 – Déroulement d'un essai (MRT informatisé).

20 secondes en condition *standard* : « *Le test démarrera dans 20 secondes* », et dans la condition *falsification* : « *Le test démarrera dans 20 secondes. Pour information, sachez que les femmes et les hommes obtiennent les mêmes résultats à ce test, ils réalisent des performances identiques.* »

Le test de RM en lui même était constitué de huit blocs de 30 essais (240 essais au total), soit dans chaque blocs 15 problèmes concernant des paires d'objets identiques et 15 différentes, avec dans les deux cas cinq problèmes pour chacun des trois angles. Les problèmes étaient présentés dans un ordre pseudo-aléatoire : les sujets ne devaient jamais produire la même réponse plus de trois fois d'affilée. Après avoir répondu à ce test, un écran invitait les participantes à prévenir l'expérimentateur qui s'assurait alors qu'elles n'avaient rencontré aucune difficulté particulière durant la passation, et qu'elles n'avaient jamais été confrontées auparavant à des tests de RM.



FIGURE 5.4 — Affichage de l'écran sensé activer l'identité de sexe des participantes. Il était demandé aux sujets de sélectionner (avec la souris) le symbole correspondant à leur sexe.

5.4.3 Attentes

Taux de réponses correctes. Premièrement, concernant le taux de réponses correctes nous n'attendons pas particulièrement d'influence des instructions, du fait même de la version du test de RM employée (présentation par paire, cf. Peters & Battista, 2008), et parce que les effets de la charge cognitive semblent essentiellement localisés au niveau des TR chez le jeune adulte (Richardson & Vecchi, 2002). Tout au plus attendons-nous un effet — très modéré — de l'amplitude des angles, car plus cette dernière est élevée et plus le problème est difficile, d'où un taux de réponses correctes éventuellement inférieur pour les angles les plus grands.

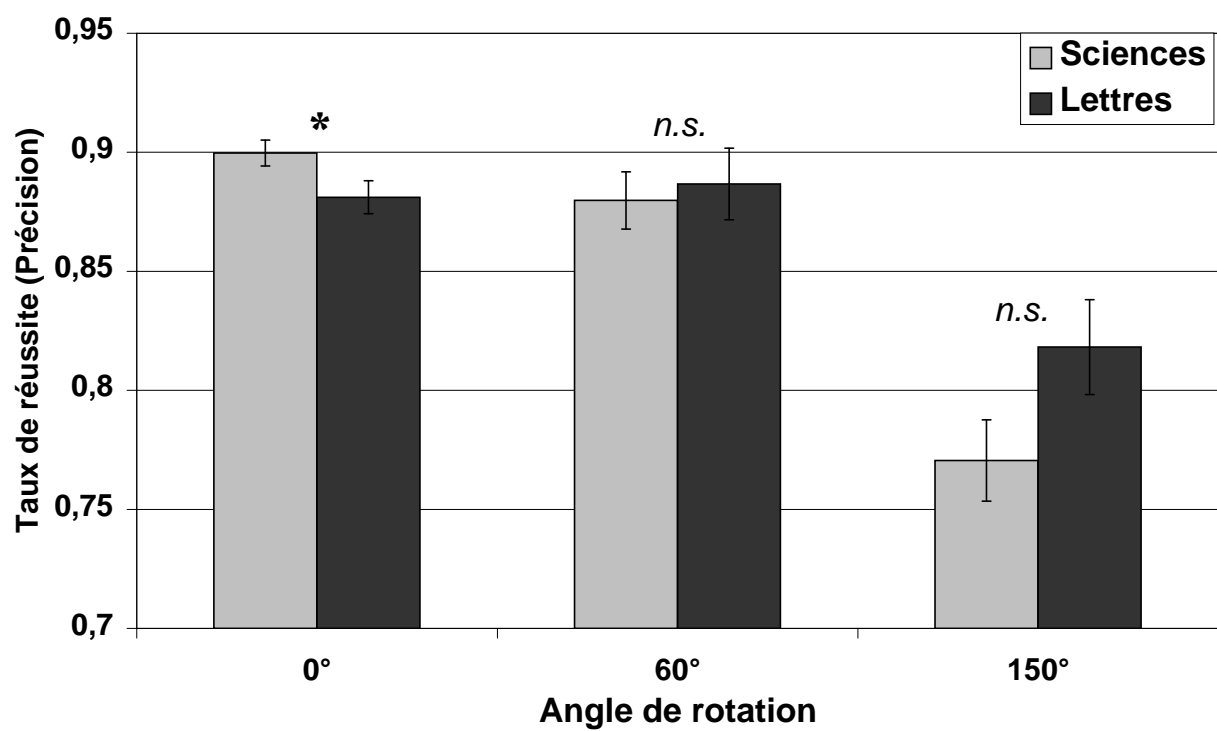
Temps de réponse. Nous attendons des TR globalement plus longs dans la condition standard (propice à l'activation implicite du stéréotype) que dans la condition de falsification, en particulier pour les étudiantes avec un bac scientifique — en principe les plus identifiées au domaine visuo-spatial. Cette attente se traduit donc par une interaction $Instruction \times Coursus$. Nous attendons surtout une interaction d'ordre deux $Instruction \times Coursus \times Angles$, l'effet du stéréotype étant en principe plus fort pour les angles les plus grands (problèmes plus difficiles).

Analyse distributionnelle. En référence au modèle de Sanders (1983), nous attendons un allongement des TR les plus longs dans la condition où le stéréotype est activé par rapport à la condition où il est falsifié, et aucune différence pour les TR les plus courts entre les deux conditions. La méthode classique d'analyse des moyennes des TR n'étant pas adapté pour vérifier cette attente, nous l'examinerons via une analyse distributionnelle des TR.

5.5 Résultats

5.5.1 Taux de réponses correctes

Une ANOVA mixte *Instructions* \times *Cursus* \times *Angles* répétée sur le dernier facteur ne montre aucun effet significatif des facteurs inter-sujets, ni des *instructions*, $F(2,71) = 1.73$, $p = .19$, ni du *cursus*, $F < 1$, ni encore d'interaction $F < 1$, ce qui est conforme aux résultats attendus pour ce type de test. Nous observons un effet classique des *angles*, $F(2,71) = 50.83$, $p < .001$; $\eta_p^2 = .41$, la performance étant nettement moins bonne pour tous les groupes lorsque l'angle est de 150 degrés ($M = .79$, $SD = .12$), par rapport aux autres angles (0 degré : $M = .90$, $SD = .04$, et 60 degrés : $M = .88$, $SD = .08$). Une interaction *Angle* \times *Cursus*, $F(2,71) = 7.87$, $p < .01$; $\eta_p^2 = .097$, indique, en l'absence de rotation, une très légère supériorité des participantes *scientifiques* ($M = .90$, $SD = .04$) sur les étudiantes *littéraires* ($M = .88$, $SD = .04$; fig. 5.5). Comme le montre la figure 5.5 ci-après, les étudiantes de lettres semblent en revanche supérieures aux étudiantes de sciences lorsque l'amplitude de la rotation à effectuer atteint 150 degrés. Mais cette différence, de fait plus grande que la précédente, n'est pas significative en raison probablement d'une variabilité plus importante dans cette modalité (150 degrés) relativement à la modalité 0 degré. L'interaction *Instructions* \times *Cursus* \times *Angles* n'est pas significative, $F(2,71) = 1.25$, $p = .28$.



* : $p < .05$

FIGURE 5.5 – Précision (pourcentage réponse correcte) - Interaction *Cursus* \times *Angles*.

5.5.2 Temps de réponse

Prétraitement des temps de réponse. Concernant l'analyse des TR, il est d'usage de supprimer les TR situés au delà de deux écart-types de la moyenne (dans un sens comme dans l'autre), car cela permet de supprimer les TR aberrants tout en permettant d'augmenter la puissance des ANOVA (Ratcliff, 1993). L'intérêt de ce traitement est évident si l'on garde à l'esprit que la validité des ANOVA repose sur une hypothèse de normalité des distributions. Or nous comptons procéder à une analyse distributionnelle qui repose sur d'autres hypothèses concernant la forme des distributions de TR, notamment leur asymétrie. Une autre procédure proposée par Ratcliff (1993) pour éliminer les valeurs aberrantes consiste à supprimer les TR les plus longs seulement (qui révèlent une trop longue hésitation qui a pu conduire le sujet à répondre au hasard par exemple), soit 5% des TR dans chaque groupe. Par ailleurs nous avons supprimé les TR trop courts (<300 ms) pour témoigner raisonnablement d'un traitement contrôlé des stimuli. Nous avons inclus dans l'analyse les réponses correctes et incorrectes, afin de conserver le plus de mesures possibles en vue de l'analyse distributionnelle.

Analyse des temps de réponse. Une ANOVA mixte *Instructions* \times *Cursus* \times *Angles* répétée sur le dernier facteur montre un effet du *cursus*, $F(1,71) = 18.72$, $p < .001$; $\eta_p^2 = .20$, les étudiantes en *sciences* étant globalement plus rapides ($M = 2418$ ms, $SD = 889$) que les étudiantes de *lettres* ($M = 3302$ ms, $SD = 888$). Nous observons également le classique effet des *angles*, $F(2,71) = 319.61$, $p < .001$; $\eta_p^2 = .81$, les TR augmentant proportionnellement à l'amplitude de l'angle (0 degré : $M = 1857$ ms, $SD = 652$; 60 degrés : $M = 2855$ ms, $SD = 1056$; 150 degrés : $M = 3573$ ms, $SD = 1437$, cf. fig. 5.6).

L'interaction *Cursus* \times *Angles* est significative, $F(2,71) = 19.58$, $p < .001$; $\eta_p^2 = .21$, l'augmentation des temps de réponse avec l'amplitude des angles de rotation étant plus marquée chez les étudiantes de lettres. Ces dernières semblent donc avoir plus de difficulté que les étudiantes de sciences lorsque la rotation mentale à effectuer est plus importante.

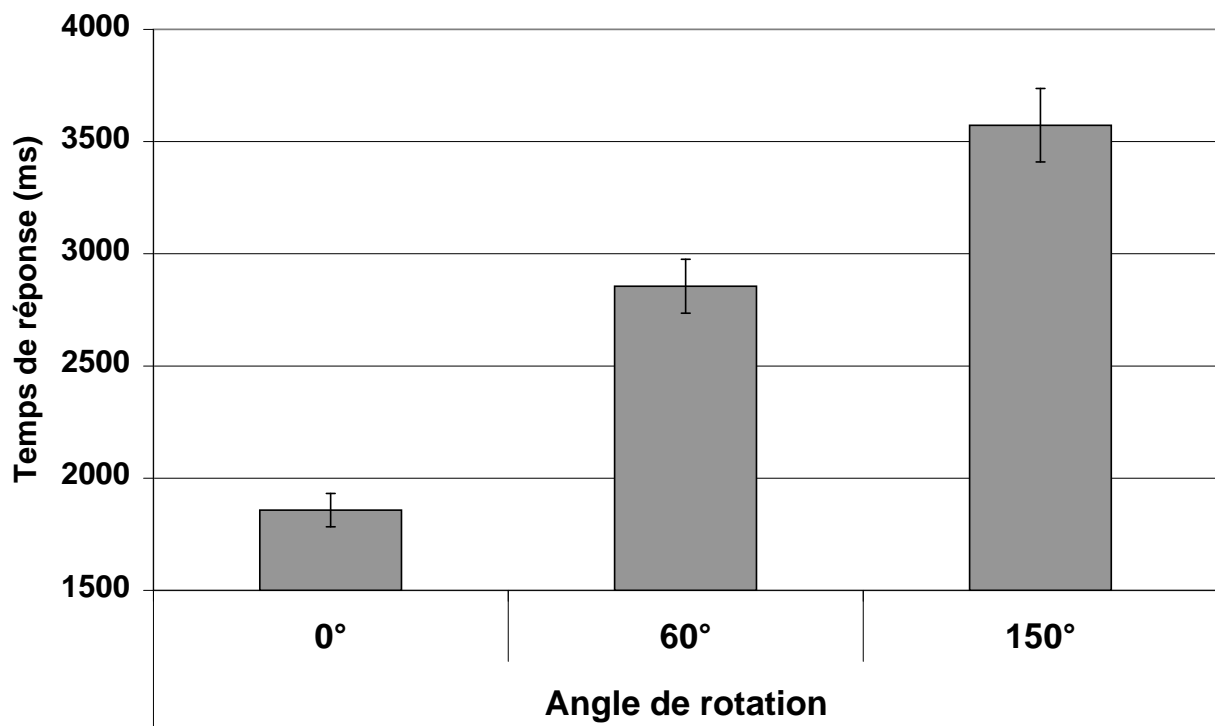
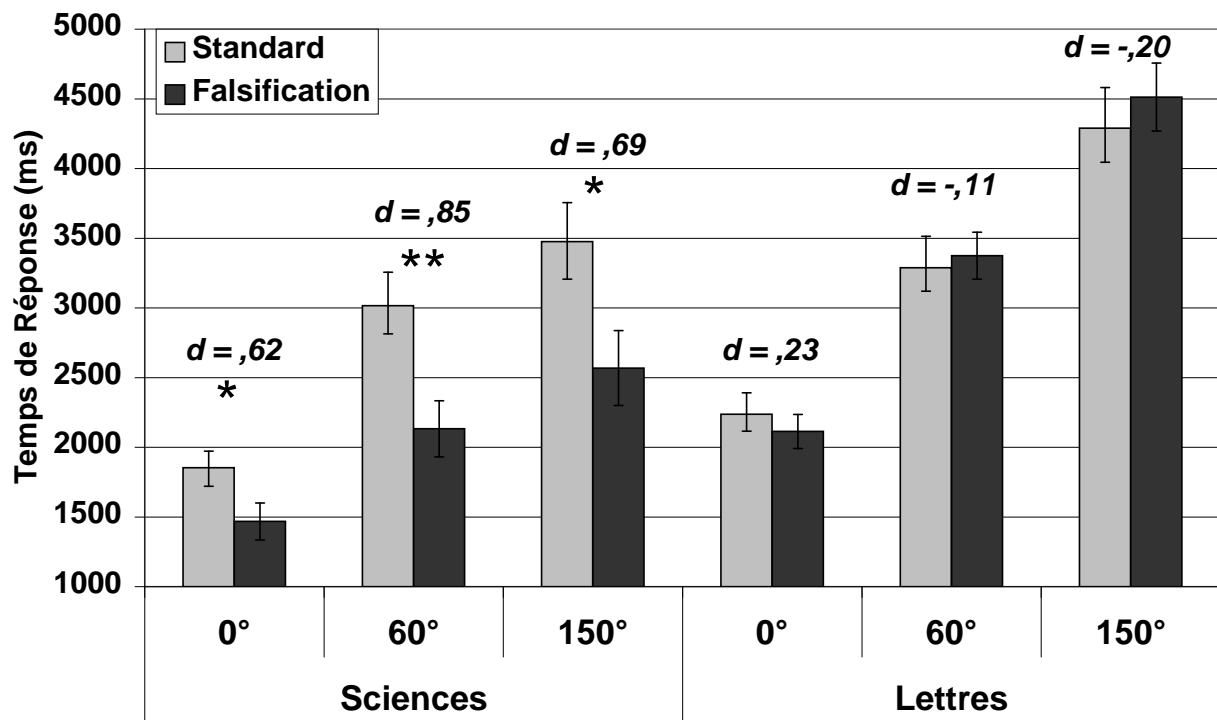


FIGURE 5.6 – Temps de Réponse - Effet de l'amplitude de l'angle.

Mais surtout, conformément à notre hypothèse centrale, nous observons bien une interaction d'ordre deux *Instructions* \times *Cursus* \times *Angles*, $F(2,71) = 5.28$, $p < .01$; $\eta_p^2 = .067$. Dans la condition de *falsification*, les étudiantes de sciences répondent significativement plus vite que tous les autres groupes, y compris les étudiantes de sciences en condition *standard*, comme l'indiquent les d de Cohen calculés entre les deux conditions dans chaque cursus (cf. fig. 5.7 p. 163). Cette plus grande rapidité s'observe surtout lorsqu'une rotation est nécessaire (angles de 60 et 150 degrés) mais survient déjà, dans une proportion moindre, lors des essais où les objets sont comparés sans devoir effectuer de RM (0 degré). Ces différences s'avèrent spécifiques aux étudiantes de sciences, ce qui est cohérent avec la littérature sur l'interférence du stéréotype.

L'effet facilitateur de la falsification chez les étudiantes de sciences dès la modalité 0 degré pose question quant aux processus impliqués. La RM semble être un processus analogique, dont l'efficacité dépend essentiellement de la vitesse de traitement de l'information visuo-spatiale. Si le *stereotype threat* induisait une baisse d'efficacité du traitement



* : $p < .05$; ** : $p < .01$ (t -tests)

FIGURE 5.7 — Temps de Réponse - Interaction *Instruction* \times *Cursus* \times *Angle*.

visuo-spatial, alors la taille de l'effet observée devrait être très réduite, sinon nulle, dès lors qu'il n'est pas nécessaire de procéder à une RM, à moins de considérer que la comparaison de figures 3D sans rotation est déjà une opération complexe, ce que nous ne pensons pas. À 0 degré, nous observons déjà une taille d'effet notable chez les étudiantes en sciences, $d = .62$, en condition de falsification celles-ci répondent en moyenne 385 ms plus vite que leurs homologues en condition *standard*. De plus, si ce sont les capacités de traitement qui sont affectés par le stéréotype, l'effet devrait être de plus en plus important lorsque l'amplitude des angles de rotations augmente. Or nous observons une taille d'effet de même ordre de grandeur pour des rotations d'amplitudes différentes — $d = .85$ à 60 degrés (écart moyen de 882 ms) et $d = .69$ à 150 degrés (écart moyen de 907 ms), l'effet semble d'ailleurs légèrement plus important à 60 degrés du fait d'une moindre variabilité dans les TR.

Ces effets, que nous prêtons à l'intervention du stéréotype, ne semblent donc pas proportionnels aux angles de rotation. Pourtant nous observons bien un effet de l'angle sur

les TR — preuve que plus l’amplitude est importante, plus le traitement est long. Donc si ce traitement est affecté par le stéréotype, alors la durée du traitement devrait s’allonger proportionnellement, c’est à dire d’autant plus que l’amplitude est importante. Cela ne correspond pas à nos observations, nous pensons donc que l’effet positif de la falsification survient au niveau du jugement lui-même, et non au niveau de l’efficacité de la RM. Selon le modèle de (Sanders, 1983), les processus énergétiques plus particulièrement impliqués à ce niveau sont l’effort et un processus de contrôle de l’effort, suite à l’évaluation de la situation. Dans un contexte de *stereotype threat*, la crainte d’échouer peut biaiser le processus d’évaluation et conduire les sujets sensibles à l’effet du stéréotype à produire un effort mal adapté (trop important) relativement à la demande de la tâche. Ce déséquilibre énergétique perturberait alors la sélection de la réponse, visible au niveau des TR. Selon Sanders, la signature de ce déséquilibre consisterait en une plus grande variabilité des TR, essentiellement visible au niveau des TR les plus longs — on constaterait alors un allongement de ces derniers — alors que la variabilité des TR rapides est difficilement détectable. Autrement dit, la perturbation des processus énergétiques par interférence du stéréotype affecterait surtout la *forme* de la fonction de distribution des TR (un allongement de la « queue » de la distribution), ce qu’une analyse classique portant sur un indice de tendance centrale comme la moyenne des TR ne permet pas de déceler. Nous proposons donc de procéder à une analyse précise de la distribution des TR dans chaque modalité.

5.5.3 Analyse distributionnelle

Les méthodes classiques d’analyse reposent sur le calcul d’estimateurs — généralement des moments de différents ordres (moyenne, variance, etc.) — à partir des valeurs mesurées sur un échantillon, pour établir les caractéristiques d’une population d’où est tiré cet échantillon. L’analyse distributionnelle adopte une perspective très différente, puisqu’il s’agit de considérer — par hypothèse — une distribution statistique particulière sensée

décrire les valeurs que prend une variable dans la population considérée, et d'ajuster les paramètres de cette distribution à partir des mesures recueillies sur l'échantillon. Au prix d'une hypothèse sur la distribution et des calculs intensifs sur un grand nombre de données que nécessite l'ajustement, l'on obtient une description relativement précise de la variable étudiée. Cette méthode d'analyse est particulièrement bien adaptée à l'étude des TR. D'une part car une abondante littérature décrit les distributions de TR envisageables (Weibull, log-normale, exponentielle, ex-gaussienne, etc., cf. Luce, 1986), d'autre part parce que les TR peuvent être recueillis aisément en quantité suffisante pour permettre des ajustements avec une erreur raisonnable. Enfin les algorithmes d'optimisation pour la plupart des ajustements sont aujourd'hui bien connus et facilement disponibles (Cousineau, Brown, & Heathcote, 2004 ; Lacouture & Cousineau, 2008 ; Van Zandt, 2000).

Nous avons choisi pour l'analyse des TR d'ajuster une distribution ex-Gaussienne, souvent choisie pour ce genre d'analyse (Heathcote et coll., 1991 ; Lacouture & Cousineau, 2008). En effet, l'ex-Gaussienne présente plusieurs caractéristiques intéressantes : ses paramètres sont plus faciles à ajuster que ceux d'autres distributions (p. ex. Weibull) qui semblent pourtant souvent décrire un peu mieux les distributions de TR (Cousineau, 1999). Mais surtout, l'interprétation des paramètres est plus aisée dans le cas de l'ex-Gaussienne que pour les autres distributions. La distribution ex-Gaussienne résulte en fait de la convolution d'une distribution Gaussienne (normale) et d'une distribution exponentielle. Elle est ainsi décrite par trois paramètres, μ , σ — liés à la composante normale — et τ , lié à la composante exponentielle (cf. fig. 5.8 p. 166). On considère généralement que la distribution ex-Gaussienne décrit des TR typiquement produits par le couplage de différents processus. Ainsi, la composante normale est parfois interprétée comme relevant des processus de nature perceptrice ou motrice, tandis que la composante exponentielle concernerait plutôt les processus de prise de décision. Par ailleurs, chaque paramètre décrit un aspect formel particulier de la distribution : μ correspond à la position de la distribution — une modification de μ entraîne une translation de la distribution —, σ détermine la forme générale de la distribution (son « étalement »), et enfin τ décrit les TR

longs (« queue » de la distribution). C'est cette dernière signification des paramètres, très descriptive, qui nous intéresse ici.

L'intérêt d'une analyse sur la base d'une ex-Gaussienne dans le cadre théorique proposé par Sanders (1983) apparaît presque immédiatement : si les mécanismes énergétiques affectent les TR longs, alors une analyse distributionnelle peut en rendre compte grâce à l'estimation des *variations spécifiques de τ* . À l'inverse, si ce sont des mécanismes computationnels qui sont en cause, la théorie prévoit en conséquence une altération systématiques des TR, c'est à dire une translation de la distribution des TR, ce qui correspondrait alors à une *variation spécifique de μ* . Ainsi notre première hypothèse, selon laquelle les étudiantes scientifiques devraient globalement mieux réussir que les étudiantes en lettres, devrait se traduire par une valeur de μ plus élevée pour ces dernières (TR globalement plus longs). Mais surtout, notre seconde hypothèse, à savoir que le phénomène de *stereotype threat* affecte essentiellement les TR longs, devrait se traduire par l'existence d'une interaction *Cursus \times Instructions*, avec un *tau* inférieur pour les étudiantes scientifiques dans la condition de *falsification* relativement aux autres groupes.

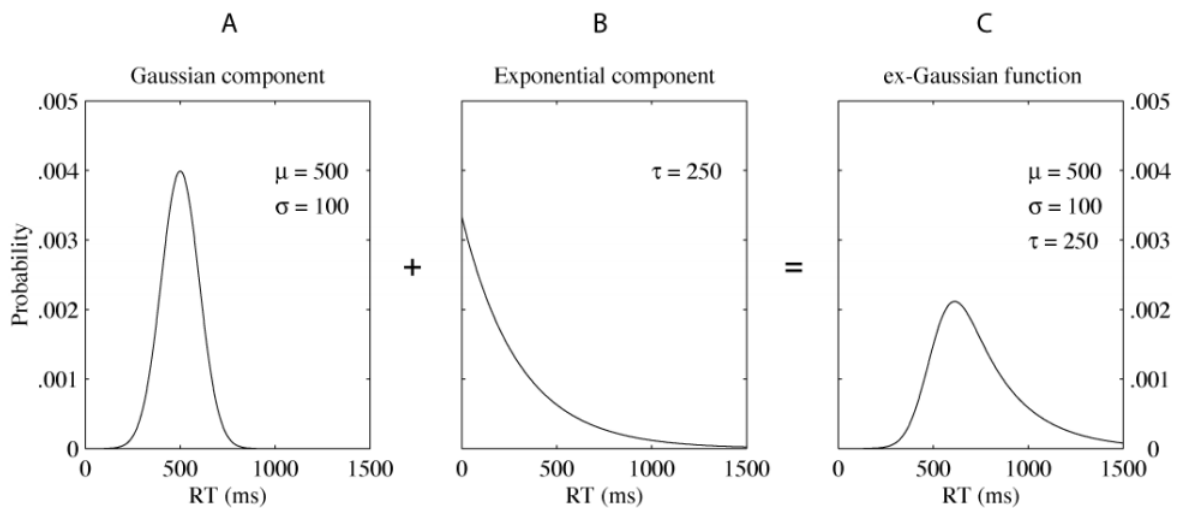


FIGURE 5.8 — La convolution d'une distribution normale avec les paramètres $\mu = 500$ et $\sigma = 100$ (A) avec une distribution exponentielle avec le paramètre $\tau = 250$ (B) résulte en une distribution ex-Gaussienne (C) (exemple tiré de Lacouture et Cousineau, 2008).

Nous avons réalisé l'estimation des paramètres à l'aide du logiciel de Cousineau et Larochelle (1997). Les TR ont été pré-traités selon la procédure indiquée section 5.5.2 (p. 161). Les paramètres des distributions ont été estimés pour chaque sujet, modalité par modalité, car les distributions doivent être construites et analysées à partir de situations strictement identiques qui induisent *a priori* des traitements cognitifs identiques. Nous avons donc considérés les distributions de TR non seulement pour chaque angle, mais aussi pour chaque type de réponse qui devait être donnée (*identique* ou *différent*), car la comparaison d'une paire d'objets identiques n'induit pas les mêmes traitements que la comparaison d'une paire d'objets différents, où il est bien connu que les TR sont plus longs. Les paramètres estimés pour les distributions de chaque sujet dans chacune de ces modalités ont par la suite été introduits dans une ANOVA *Cursus* \times *Instructions*. Du fait du grand nombre d'analyses résultantes (trois paramètres sont estimés pour chacune des six modalités expérimentales) nous présentons les résultats sous forme de tableau à chaque fois.

Paramètre τ . L'interaction attendue *Cursus* \times *Instructions* est significative pour toutes les modalités sauf *identique / 60 degrés*, où elle n'est que tendancielle, $F(1,73) = 3.77$, $p = .056$ (cf. tableau 5.1 p. 168). Des contrastes orthogonaux opposant les groupes selon le cursus et les instructions montrent que τ est significativement inférieur chez les étudiantes de sciences en condition de *falsification* relativement à leurs homologues en condition *standard* et aux étudiantes de lettres en condition de *falsification*. Les TR les plus longs des étudiantes en sciences en condition de *falsification* sont donc inférieurs aux TR les plus longs de tous les autres groupes. La figure 5.9 p. 169 donne une représentation graphique de ces interactions.

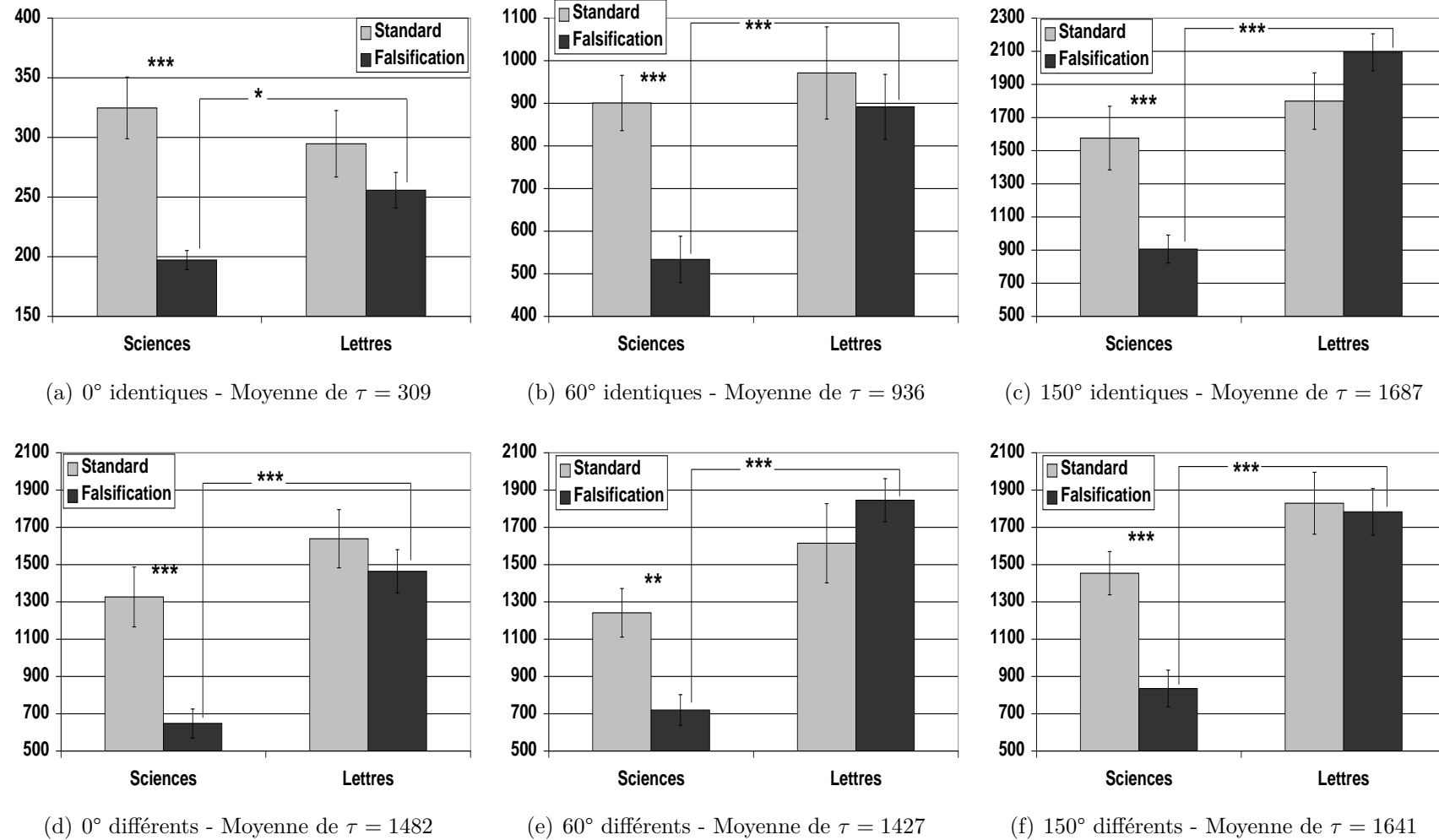
Nous observons également un autre phénomène : les valeurs de τ sont aussi affectées par le type de comparaison demandé (objets identiques vs. différents). Lors de la comparaison de paire d'objets identiques, les ordres de grandeur des paramètres augmentent de manière notable lorsque l'amplitude des angles augmente. Par exemple, globalement, pour

la modalité *identique / 0 degré*, la valeur moyenne est $\tau = 309$ (cf. fig. 5.9a²), $\tau = 936$ dans la modalité *identique / 60 degrés* (cf. fig. 5.9b), et finalement $\tau = 1687$ dans la modalité *identique / 150 degrés* (cf. fig. 5.9c). Or, lorsque les comparaisons portent sur des objets différents, pour les trois angles, la valeur de τ est du même ordre de grandeur, autour de 1500 (cf. fig. 5.9 d, e et f). Ainsi, une ANOVA mixte *Instructions* \times *Cursus* \times *Angle* répétée sur le dernier facteur, montre une interaction d'ordre deux significative pour les paires d'objets identiques, $F(2,71) = 6.77$, $p < .005$; $\eta_p^2 = .085$, alors que cette interaction n'est pas significative pour les paires d'objets différents, $F < 1$. Il semble donc que lorsque le sujet doit répondre négativement (« objets différents »), l'amplitude de l'angle ne détermine pas à lui seul le processus d'évaluation et de sélection de la réponse. En terme de durée de traitement, la durée de la comparaison et de la décision semble alors prendre le pas sur celle de la RM proprement dite.

TABLE 5.1 – Analyses (ANOVA) pour le paramètre τ , estimé pour chaque type de paire d'objets (*Identiques* vs. *Différents*). Les effets statistiquement significatifs sont en gras.

Identiques	0°	<i>Cursus</i> , $F < 1$ <i>Instructions</i> , $F(1,73) = 19.98$, $p < .001$, $\eta_p^2 = .22$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 5.66$, $p < .05$, $\eta_p^2 = .072$
	60°	<i>Cursus</i> , $F(1,73) = 8.39$, $p < .005$, $\eta_p^2 = .10$ <i>Instructions</i> , $F(1,73) = 9.12$, $p < .005$, $\eta_p^2 = .11$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 3.77$, $p = .056$, $\eta_p^2 = .049$
	150°	<i>Cursus</i> , $F(1,73) = 26.15$, $p < .001$, $\eta_p^2 = .26$ <i>Instructions</i> , $F(1,73) = 1.85$, $p = .178$, $\eta_p^2 = .025$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 12.22$, $p < .001$, $\eta_p^2 = .14$
Différents	0°	<i>Cursus</i> , $F(1,73) = 20.52$, $p < .001$, $\eta_p^2 = .22$ <i>Instructions</i> , $F(1,73) = 11.73$, $p < .001$, $\eta_p^2 = .14$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 4.09$, $p < .05$, $\eta_p^2 = .053$
	60°	<i>Cursus</i> , $F(1,73) = 33.64$, $p < .001$, $\eta_p^2 = .32$ <i>Instructions</i> , $F(1,73) = 1.26$, $p = .265$, $\eta_p^2 = .017$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 8.48$, $p < .005$, $\eta_p^2 = .104$
	150°	<i>Cursus</i> , $F(1,73) = 28.22$, $p < .001$, $\eta_p^2 = .28$ <i>Instructions</i> , $F(1,73) = 7.12$, $p < .01$, $\eta_p^2 = .089$ <i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 5.28$, $p < .05$, $\eta_p^2 = .067$

2. Pour faciliter la lisibilité de l'interaction dans chaque modalité, nous avons adapté l'échelle de chaque graphique de la figure 5.9 à l'ordre de grandeur de la moyenne de τ .



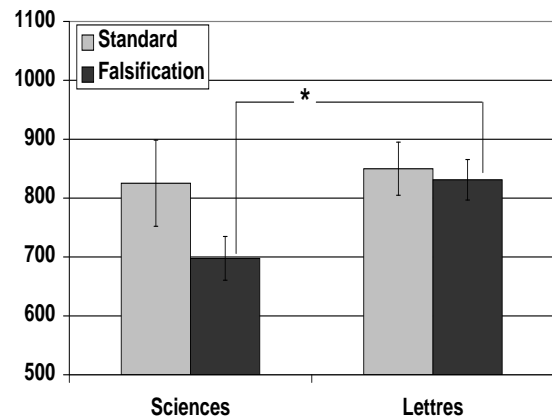
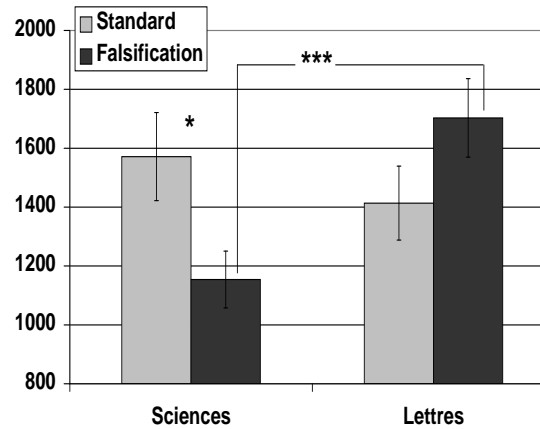
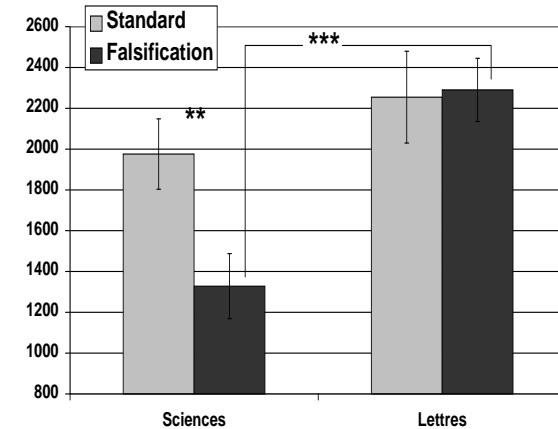
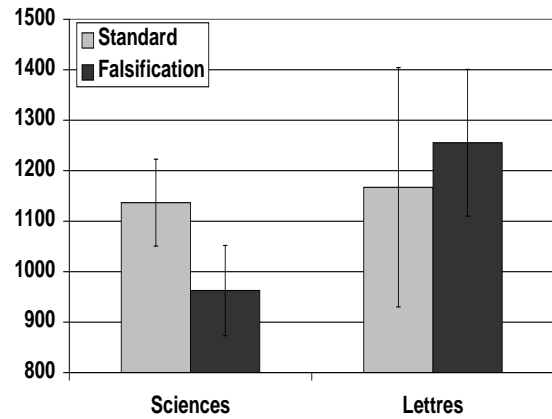
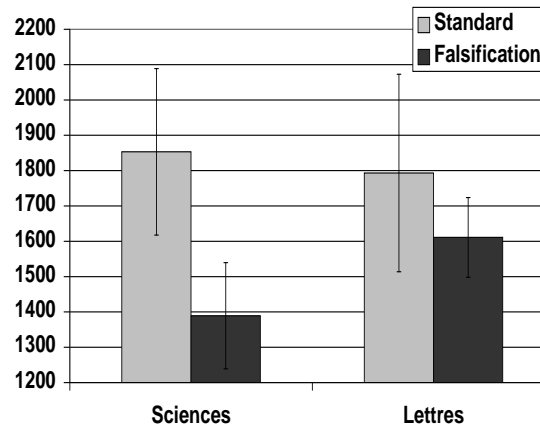
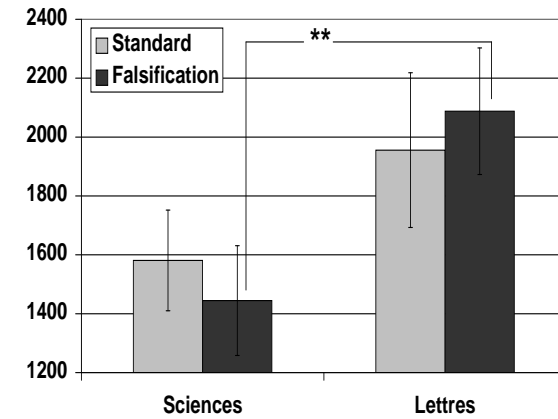
* : $p < .05$; ** : $p < .01$; *** : $p < .001$

FIGURE 5.9 – Paramètre τ , interactions *Cursus* \times *Instructions* pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de τ dans chaque modalité.

Paramètre μ . Nous attendions un effet du cursus sur μ , sensé montrer la supériorité des étudiantes de sciences en RM par des réponses globalement plus rapides, nous n’observons cet effet seulement pour les angles de 150 degrés (cf. tab. 5.2 ci-dessous). En outre, l’analyse révèle une interaction *Cursus* \times *Instructions* significative dans la modalité *identique* / 60 degrés (celle pour laquelle l’interaction concernant τ était juste tendancielle). Les contrastes orthogonaux opposant les groupes selon le cursus et les instructions montrent que la valeur de μ est bien inférieure pour le groupe *science* / *falsification* relativement aux groupes *science* / *standard* et *lettre* / *falsification* (cf. fig. 5.10 p. 171). Nous constatons également que l’amplitude des angles affecte différemment la valeur de μ selon le type de comparaison. L’ANOVA mixte *Instructions* \times *Cursus* \times *Angle* répétée sur le dernier facteur montre une interaction d’ordre deux tendancielle pour les paires d’objets identiques, $F(2,71) = 3.24$, $p = .051$; $\eta_p^2 = .085$. Cette interaction n’est pas significative pour les paires d’objets identiques, $F < 1$.

TABLE 5.2 – Analyses (ANOVA) pour le paramètre μ , estimé pour chaque type de paire d’objets (*Identiques* vs. *Différents*). Les effets statistiquement significatifs sont en gras.

Identiques	0°	<i>Cursus</i> , $F(1,73) = 2.53$, $p = .116$, $\eta_p^2 = .033$
		<i>Instructions</i> , $F(1,73) = 2.17$, $p = .145$, $\eta_p^2 = .029$
		<i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 1.20$, $p = .277$, $\eta_p^2 = .016$
	60°	<i>Cursus</i> , $F(1,73) = 2.34$, $p = .130$, $\eta_p^2 = .031$
		<i>Instructions</i> , $F < 1$
		<i>Instructions</i> \times <i>Cursus</i>, $F(1,73) = 7.65$, $p < .01$, $\eta_p^2 = .095$
	150°	<i>Cursus</i>, $F(1,73) = 11.99$, $p < .001$, $\eta_p^2 = .14$
		<i>Instructions</i> , $F(1,73) = 2.91$, $p = .092$, $\eta_p^2 = .038$
		<i>Instructions</i> \times <i>Cursus</i> , $F(1,73) = 3.63$, $p < .061$, $\eta_p^2 = .047$
Différents	0°	<i>Cursus</i> , $F(1,73) = 1.40$, $p = .240$, $\eta_p^2 = .019$
		<i>Instructions</i> , $F < 1$
		<i>Instructions</i> \times <i>Cursus</i> , $F < 1$
	60°	<i>Cursus</i> , $F < 1$
		<i>Instructions</i> , $F(1,73) = 2.78$, $p = .099$, $\eta_p^2 = .037$
		<i>Instructions</i> \times <i>Cursus</i> , $F < 1$
	150°	<i>Cursus</i>, $F(1,73) = 5.79$, $p < .05$, $\eta_p^2 = .073$
		<i>Instructions</i> , $F < 1$
		<i>Instructions</i> \times <i>Cursus</i> , $F < 1$

(a) 0° identiques - Moyenne de $\mu = 801$ (b) 60° identiques - Moyenne de $\mu = 1492$ (c) 150° identiques - Moyenne de $\mu = 2115$ (d) 0° différents - Moyenne de $\mu = 1151$ (e) 60° différents - Moyenne de $\mu = 1823$ (f) 150° différents - Moyenne de $\mu = 1768$

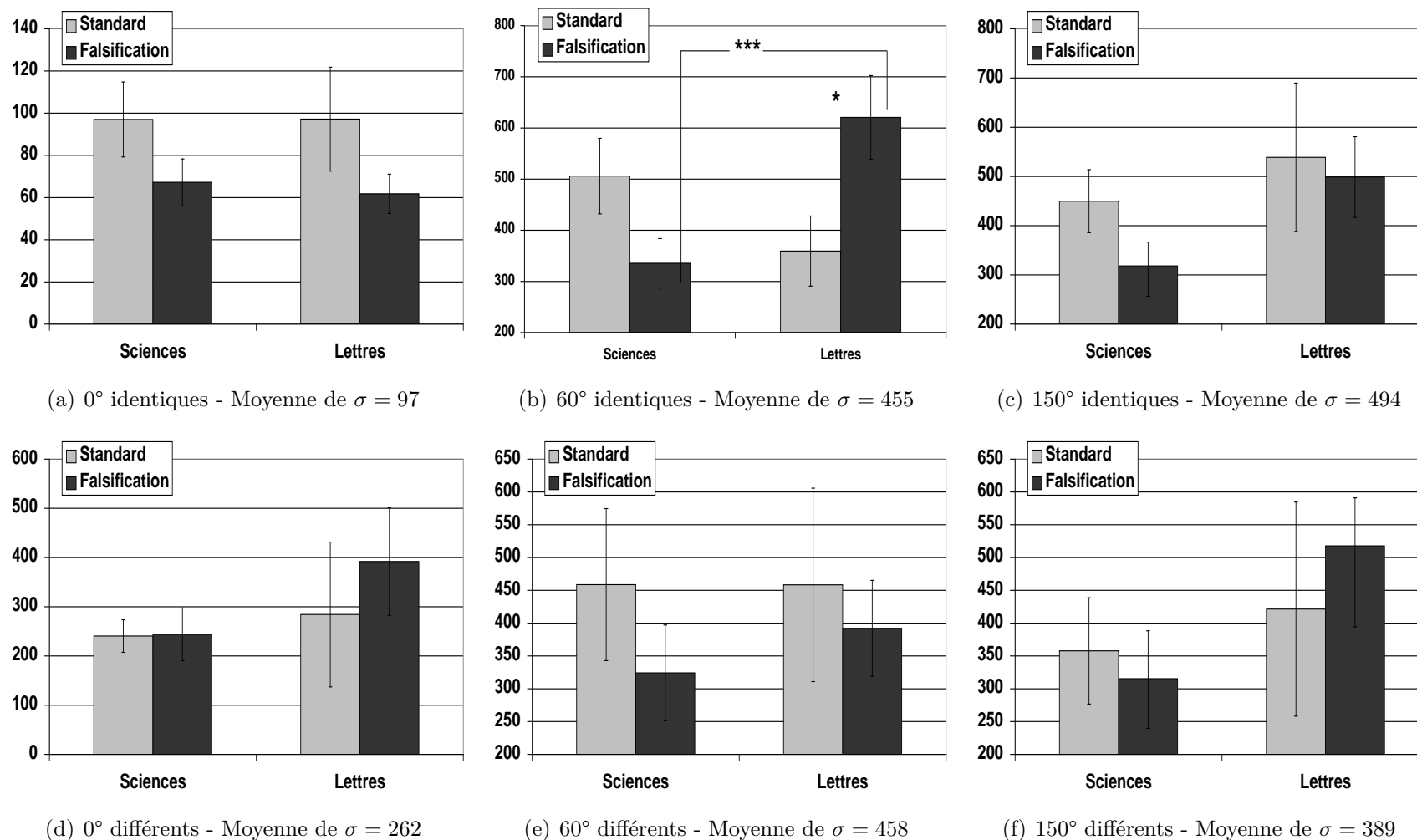
*** : $p < .001$; ** : $p < .01$; * : $p < .05$

FIGURE 5.10 – Paramètre μ , interactions *Cursus* \times *Instructions* pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de μ dans chaque modalité.

Paramètre σ . Concernant le paramètre σ , outre un effet principal des instructions dans la modalité *identique/0 degré*, les analyses révèlent encore une fois une interaction $Cursus \times Instructions$ dans la modalité *identique/60 degrés* (cf. tab. 5.3 p. 173). Rappelons ici que pour cette dernière modalité, cette interaction était également significative pour τ et σ . Il est possible que l'ajustement des paramètres dans cette modalité n'ait pas abouti à des paramètres aussi indépendants que ce qui est prévu par le modèle, peut-être du fait du nombre limité de TR recueillis, ce qui a pu peser sur la qualité de l'ajustement. Malgré cela, nous observons bien que les facteurs expérimentaux n'ont que peu d'effets sur σ alors que ces effets sont massivement localisés au niveau du paramètre τ , comme attendu. Par ailleurs, nous constatons encore que l'amplitude des angles affecte différemment la valeur de σ selon le type de paire présentée. L'ANOVA mixte $Instructions \times Cursus \times Angle$ répétée sur le dernier facteur montre une interaction d'ordre deux significative pour les paires d'objets identiques, $F(2,71) = 3.70$, $p < .05$; $\eta_p^2 = .048$. Cette interaction n'est pas significative pour les paires d'objets identiques, $F < 1$.

TABLE 5.3 – Analyses (ANOVA) pour le paramètre σ , estimé pour chaque type de paire d'objets (Identiques vs. Différents). Les effets statistiquement significatifs sont en gras.

Identiques	0°	<i>Cursus</i> , $F < 1$ <i>Instructions</i>, $F(1,73) = 4.50$, $p < .05$, $\eta_p^2 = .058$ <i>Instructions</i> \times <i>Cursus</i> , $F < 1$
		60° <i>Cursus</i> , $F(1,73) = 1.01$, $p = .317$, $\eta_p^2 = .014$ <i>Instructions</i> , $F < 1$ <i>Instructions</i> \times <i>Cursus</i>, $F(1,73) = 9.86$, $p < .005$, $\eta_p^2 = .12$
	150°	<i>Cursus</i> , $F(1,73) = 2.40$, $p = .126$, $\eta_p^2 = .032$ <i>Instructions</i> , $F < 1$ <i>Instructions</i> \times <i>Cursus</i> , $F < 1$
Différents	0°	<i>Cursus</i> , $F(1,73) = 1.20$, $p = .276$, $\eta_p^2 = .016$ <i>Instructions</i> , $F < 1$ <i>Instructions</i> \times <i>Cursus</i> , $F < 1$
		60° <i>Cursus</i> , $F < 1$ <i>Instructions</i> , $F(1,73) = 1.34$, $p = .252$, $\eta_p^2 = .018$ <i>Instructions</i> \times <i>Cursus</i> , $F < 1$
	150°	<i>Cursus</i> , $F(1,73) = 1.50$, $p = .224$, $\eta_p^2 = .020$ <i>Instructions</i> , $F < 1$ <i>Instructions</i> \times <i>Cursus</i> , $F < 1$



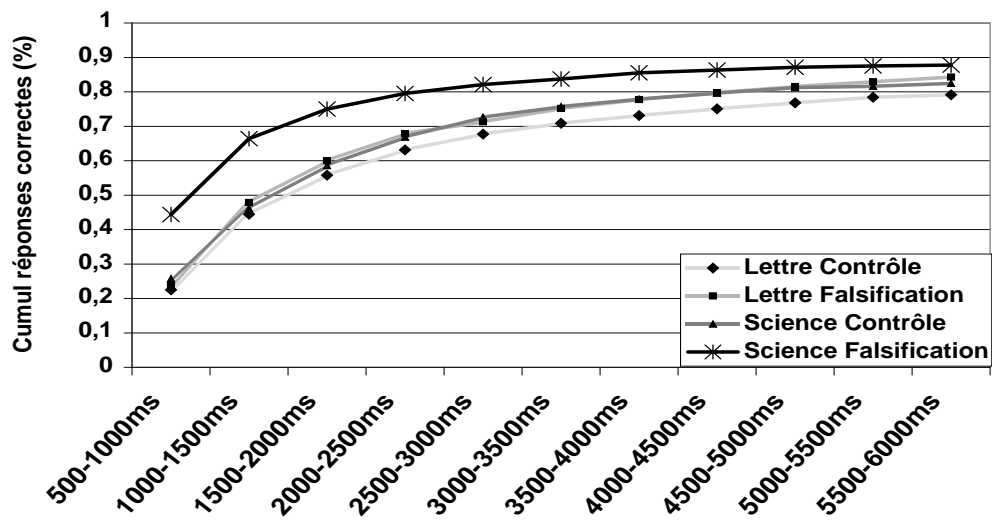
*** : $p < .001$; * : $p < .05$

FIGURE 5.11 – Paramètre σ , interactions *Cursus* \times *Instructions* pour chaque modalité (angles et type de comparaison). Les échelles des graphiques sont ajustées à la valeur moyenne globale de σ dans chaque modalité.

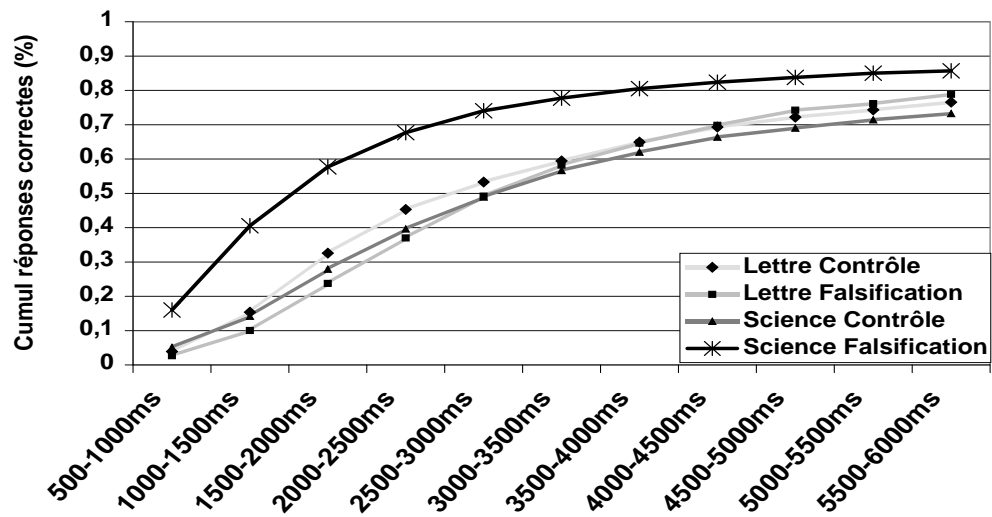
En conclusion, la falsification a un effet positif sur la vitesse des réponses des étudiantes en sciences. Cet effet semble être dû au processus de prise de décision liés à des mécanismes énergétiques, car le gain en rapidité s'observe essentiellement au niveau des TR les plus longs. D'autres résultats vont dans ce sens : l'effet facilitateur s'observe également au niveau des problèmes qui ne nécessitent aucune rotation mentale, et au niveau des problèmes où l'amplitude de l'angle de rotation ne semble pas affecter les TR (paires identiques vs. différentes).

5.5.4 Rapport vitesse-précision

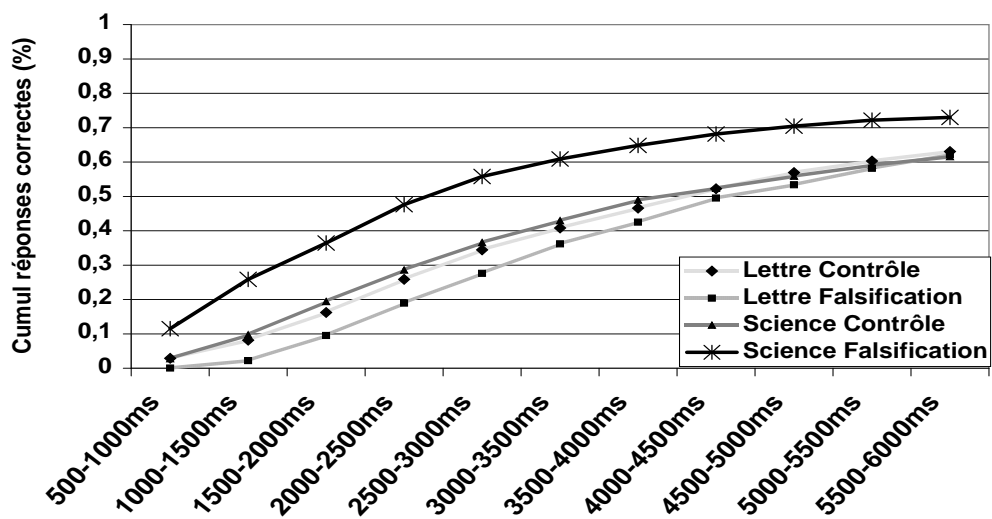
Si les étudiantes en sciences répondent plus rapidement dans la condition de falsification, il faut nous assurer que cette rapidité n'est pas associée à une baisse du nombre de réponses justes. En effet, la falsification du stéréotype pourrait simplement agir au niveau de la stratégie d'échange vitesse - précision et les pousser à répondre plus vite en faisant plus d'erreurs. L'absence d'effet des instructions sur le taux de réponses correctes (cf. section 5.5.1 p. 159) est un premier élément de réponse : les étudiantes en sciences en condition de *falsification* répondent plus vite mais pas moins juste, elles semblent donc plus efficaces. Il nous a paru intéressant de proposer en complément une lecture graphique des rapports entre vitesse et précision dans les différents groupes. Pour réaliser les graphes de la figure 5.12, nous focalisant sur les TR entre 500 ms et 6000 ms dans chaque condition, nous avons calculé pour chaque « tranche » de 500 ms le taux cumulé de réponses justes. Ceci confirme nos analyses précédentes : nous constatons que les participantes avec une formation scientifique se distinguent nettement dans la condition de *falsification*. Leurs réponses sont plus rapides pour une performance globale du même ordre que dans les autres conditions, et ce, de manière surprenante, même dans la condition de comparaison des objets sans rotation (0 degré).



(a) Pas de rotation (0°)



(b) Angle de rotation (60°)



(c) Angle de rotation (150°)

FIGURE 5.12 – Pourcentages cumulés des réponses correctes en fonction de la rapidité des réponses.

5.6 Discussion

Dans cette étude, nous nous attendions à ce que des étudiantes en sciences, susceptibles d'être plus identifiées au domaine visuo-spatial, montrent de meilleures performances en RM suite à la falsification du stéréotype relativement à leurs homologues dans une condition *standard* et à des étudiantes de lettres, moins identifiées au domaine, dans les mêmes conditions. Nos résultats sont conformes à ces attentes, nous observons bien un effet positif de la falsification sur les temps de réponse des étudiantes en sciences, plus rapides que toutes les étudiantes dans les autres conditions, sans que la justesse de leurs réponses n'en souffre. Ceci confirme donc d'une part la validité du paradigme de falsification sur une tâche informatisée de RM passée en laboratoire. D'autre part, ces résultats confirment également l'idée que toutes les femmes ne sont pas sensibles à la manipulation du stéréotype, les étudiantes de lettres, *a priori* moins performantes et moins identifiées au domaine visuo-spatial (cf. étude 3), y semblent moins sensibles.

En outre, mesurer la performance par le biais des TR a permis de mettre à jour l'effet positif de la falsification également pour des problèmes sans rotation (0 degré), donc très faciles. Ce résultat va à l'encontre des attentes classiques concernant le *stereotype threat*, sensé ne pas survenir dans le cas de problèmes faciles. L'interférence du stéréotype semble donc avoir un effet au niveau de la rapidité alors que cet effet est indétectable au niveau des performances mesurées en terme de justesse des réponses, distinction qui n'avait jamais été opérée auparavant. En effet la littérature sur le *stereotype threat* s'est surtout concentrée sur des tâches de haut niveau (test d'intelligence, de vocabulaire, de mathématiques, etc.) qui ne permettent pas un recueil précis des TR. Seules Wraga et coll. (2006) et Wraga et coll. (2007) avaient proposées à notre connaissance de recueillir des TR sur une tâche de RM peu classique où les sujets devaient imaginer une scène depuis divers points de vue et indiquer si un élément particulier était alors toujours visible ou non. Ces auteurs n'avaient alors montré aucun effet du stéréotype sur les TR, mais plutôt sur le taux d'erreur, avec un impact assez faible, de l'ordre de 6 à 8%. Nous avons préféré

une tâche de RM plus classique, du type de celle proposée initialement par Shepard et Metzler (1971), sur laquelle nous n'avons observé aucun effet sur le nombre de réponses correctes, comme c'est souvent le cas avec cette tâche (Peters & Battista, 2008). Par contre nous avons pu montrer un effet positif et massif de la falsification sur les TR. Ce résultat original ouvre un champ d'investigation important pour la compréhension du phénomène de *stereotype threat* car nous disposons aujourd'hui d'un large éventail de méthodes qui permettent d'obtenir une quantité d'information non-négligeable à partir de l'analyse des TR sur les processus cognitifs sous-jacents.

Précisément, nous avons mis en œuvre l'une de ces méthodes (l'analyse distributionnelle) en procédant à l'analyse fine de la forme des distributions de TR recueillis, pour constater que les réponses des sujets ne sont pas toutes affectées de la même manière par la manipulation du stéréotype. En effet, dans la condition de *falsification* les distributions de TR des étudiantes en sciences sont moins allongées relativement à celles des autres étudiantes dans toutes les autres conditions. Ainsi, la plus grande rapidité — observée en moyenne — de ces étudiantes semble due à un nombre moins important de TR longs. Nous interprétons ce résultat à la lumière du modèle de Sanders (1983), selon lequel une telle variabilité dans les TR longs est le signe que le phénomène capturé n'affecte pas toutes les réponses systématiquement, ce qui est caractéristique des mécanismes énergétiques (contrôle et ressources attentionnelles). Ainsi nos résultats apportent un argument en faveur de l'idée que l'intervention du stéréotype n'affecte pas la capacité de traitement en elle-même, mais les mécanismes d'évaluation et d'attribution des ressources disponibles, soit *l'effort* selon Sanders (1983), ou des processus d'auto-régulation et le contrôle de l'attention selon d'autres auteurs (Schmader et coll., 2008). Nous pensons apporter une contribution importante à la compréhension du phénomène en question, car nous avons validé une méthode qui permet de mesurer indirectement l'intervention de ces mécanismes sans avoir recourt à des mesures plus explicites ou directes (questionnaires, mesures physiologiques, etc.), souvent lourdes et qui peuvent interférer avec le test lui-même.

À partir de ce résultat, nous pouvons nous poser une question d'ordre général sur le *stereotype threat* : le phénomène surviendrait peut-être plus souvent qu'imaginé jusque ici. En fait, il serait seulement indétectable sans des mesures aussi précises que le recueil des TR. Dans le cas de problème faciles, même si une interférence du stéréotype survient, les sujets peuvent tout de même très bien réussir selon un critère de justesse des réponses (celui retenu par toutes les études jusqu'ici). Pourtant si l'on analysait la performance en terme de rapidité, celle-ci pourrait très bien être affectée, sans que cela transparaisse au niveau de la justesse des réponses, vu que les délais en question restent très courts. Mais il est envisageable qu'une pression temporelle très forte, qui en quelque sorte serait perçue comme une augmentation perceptible de la difficulté, permette de voir apparaître un effet du stéréotype sur la justesse des réponses à ces problèmes faciles, du fait de la légère baisse de rapidité qui pourrait survenir.

Cependant notre étude présente quelques limites. Concernant l'analyse distributionnelle, nous avons construit notre test de RM de manière à obtenir un nombre raisonnable de mesures dans chaque modalité. Nous avons fait face à deux contraintes contradictoires : d'une part l'estimation des paramètres nécessite un nombre important de données pour être fiable, mais d'autre part des pré-tests nous ont montré que les sujets avaient des difficultés à mener à son terme un test de RM trop long, qui par sa difficulté demande un effort important. Pour une première étude dans le genre, nous avons essayé d'équilibrer ces deux contraintes en réalisant le nombre minimum de mesures qui nous permettait de réaliser une estimation des paramètres. Néanmoins il conviendrait de confirmer nos résultats par un test où le nombre de mesures serait bien plus important, par exemple en réduisant le nombre de modalités. Par ailleurs, si nous avons observé une variation de la performance des femmes entre nos conditions, nous n'avons pas testé d'hommes, ce qui ne nous permet pas d'apprécier quelle part de la différences de sexe nous pourrions expliquer.

5.7 Conclusion

Dans cette étude, nous avons validé le paradigme de falsification sur un test informatisé de RM, ce qui nous a permis de montrer les effets du stéréotype sur les temps de réponse, ce qu'aucune étude n'avait fait jusqu'à présent à notre connaissance. Ce dispositif nous a permis d'observer un résultat inattendu qui contredit un élément fondamental de la théorie de Steele : l'effet positif de la falsification sur les TR concernant des problèmes sensés être faciles (la simple comparaison de deux objets, sans rotation) et donc insensibles à l'interférence du stéréotype. Ce résultat, associé à une analyse plus fine et originale des distributions de TR (analyse distributionnelle), suggère que ce ne sont pas les processus de traitement de l'information visuo-spatiale eux-mêmes qui subissent les effets du stéréotype, mais des processus de contrôle et de régulation « énergétiques » (selon la terminologie de Sanders, 1983).

Discussion Générale

Notre objectif dans cette thèse était de contribuer à la connaissance des mécanismes sous-jacents à l'une des différences de sexe les plus robustes en psychologie cognitive, à partir d'une confrontation de deux champs de recherche jusqu'ici peu articulés. Prenant appui sur les travaux de Steele et de ses collaborateurs (p. ex. Steele & Aronson, 1995 ; Steele, 1997), nous avons cherché à montrer le fondement psycho-social de l'infériorité supposée des femmes dans le domaine visuo-spatial.

La théorie du *stereotype threat* indique comment certains stéréotypes, une fois activés au cours du test, interfèrent avec la performance, au point de créer une réalité qui lui est conforme. Il s'agit d'une approche situationnelle qui rompt avec l'angle macro-social sous lequel les facteurs sociaux de la différence de sexe ont jusqu'ici été abordés. Le paradigme expérimental permettant de mettre en évidence la survenue de l'interférence du stéréotype consiste à comparer une condition où ce dernier est soit explicitement, soit implicitement activé, à une condition où il est explicitement falsifié (Spencer et coll., 1999). Il s'agit donc d'une intervention visant à annuler le phénomène pour le mettre en évidence, en d'autres termes, il s'agit d'une réelle manipulation expérimentale d'un facteur social, ce qui est relativement inédit dans l'étude des compétences visuo-spatiales.

Nous présentons ci-dessous nos études expérimentales qui ont permis de valider cette approche. Nous exposerons ensuite quelques éléments de réflexion sur les enjeux sociétaux de ce travail, pour conclure sur plusieurs perspectives de recherche.

Validation et intervention du stéréotype au cours de la rotation mentale

Notre première étude nous a permis de valider, sur un important échantillon de collégiens, le paradigme de *falsification* du stéréotype au cours d'un test de rotation mentale : le MRT de Vandenberg et Kuse (1978). Conformément à nos attentes, la proportion de collégiennes atteignant les plus hauts scores était bien plus importante dans la condition de falsification que dans les conditions d'activation explicite ou supposée implicite du stéréotype. La principale originalité de cette première étude a été de proposer une analyse de la distribution de performance, qui nous a permis de détecter des effets qui seraient passés inaperçus autrement.

Le regard porté sur ces différences s'est sans cesse affiné — comme les analyses permettant de les caractériser. Nous l'avons déjà évoqué dans notre revue de question sur la rotation mentale : la communauté scientifique dans ce domaine est passée d'une perception des différences de sexe assez globale (Maccoby & Jacklin, 1974), à l'idée que ces différences étaient caractérisées par différentes tailles d'effet (Hyde, 1981), en fonction d'une part des tests utilisés (Linn & Petersen, 1985 ; Voyer et coll., 1995), et d'autre part de la portion de la distribution de performance considérée (Hedges & Nowell, 1995 ; Halpern, 2000). Nous avons pleinement intégré cette évolution dans notre première étude conduite auprès d'un échantillon que nous avons voulu de grande taille. Il s'agit d'une démarche originale, en outre justifiée par la théorie du *stereotype threat* : l'effet du stéréotype est sensé être localisé chez les sujets pour qui l'issue du test représente un enjeu. Or nous l'avons vu, ce dernier existe surtout pour les sujets situés dans la partie haute de la distribution de performance et qui ont aussi le plus à perdre en cas d'échec. À notre connaissance, aucun chercheur dans le domaine du *stereotype threat* ne s'était donné les moyens de tester l'effet en question à différents niveaux de performances comme nous l'avons fait dans l'étude 1, dont les résultats constituent par conséquent une contribution véritablement originale (article en court de rédaction).

Les résultats de l'étude 1 alertent également sur le fait qu'il existe une dynamique au phénomène qui correspond à une érosion des performances au cours du test. En comparant les deux sections du test nous avons constaté que, dans les conditions *contrôle* et d'*activation* explicite du stéréotype, la performance des filles baissait, alors qu'elle se maintenait lorsque le stéréotype était falsifié (cf. aussi étude 3). Il s'agit là d'une autre contribution intéressante. Auparavant, au sein de la littérature « classique » sur les différences de sexe en RM, Peters (2005) avait déjà constaté que ces différences s'exprimaient de plus en plus clairement au fur et à mesure de la passation du test. Il jugeait néanmoins ce résultat difficile à interpréter. Nous apportons pour notre part un nouvel éclairage à ce sujet, éclairage par ailleurs compatible avec les apports de Cadinu et coll. (2005), qui constataient l'augmentation progressive du nombre de pensées négatives au cours du test proposé. Enfin, et peut-être surtout, les résultats de notre première étude contribuent à montrer la malléabilité d'une différence de sexe considérée par ailleurs comme éminemment robuste. Cette différence s'estompe en effet dès lors que les sujets bénéficient avant le test d'une information de nature à falsifier le stéréotype dont leur groupe de sexe fait l'objet.

Un autre point fort à retenir par rapport à certaines études antérieures sur l'intervention du stéréotype de genre en RM (Martens et coll., 2006 ; Moè & Pazzaglia, 2006 ; Wraga et coll., 2006, 2007) est notre choix de reprendre la proposition de falsification de Spencer et coll. (1999). En effet, dans les études citées précédemment, les participants étaient confrontés à un *feedback* de comparaison sociale leur précisant directement la supériorité d'un groupe de sexe sur l'autre, *feedback* en outre de nature explicitement stéréotypique ou contre-stéréotypique. Ce type de *feedback* complique l'interprétation puisque l'on ne sait plus si les différences observées entre les différents groupes traduisent en effet la mobilisation du stéréotype dont on fait l'hypothèse ou la tonalité émotionnelle susceptible d'accompagner tout *feedback* positif ou négatif.

En outre, la falsification offre l'avantage — paradoxal — d'activer le stéréotype car elle s'y réfère par définition explicitement. Dans le cas où cette falsification s'avère efficace, à l'instar de nos travaux, l'effet obtenu nous permet d'invalider toutes les explications du *stereotype threat* fondées ou dérivées de « l'hypothèse idéomotrice » (p. ex. « *priming* comportemental »), où l'on considère que la simple activation d'un stéréotype a pour effet de provoquer automatiquement le comportement spécifié par celui-ci (Dijksterhuis, Aarts, Bargh, & Knippenberg, 2000 ; Wheeler & Petty, 2001). La falsification, qui nécessairement active les catégories de sexe, est de fait propice à l'activation du stéréotype. Dans nos travaux comme dans d'autres (Spencer et coll., 1999 ; Désert et coll., 2002), elle supprime néanmoins les différences invoquées par le stéréotype.

Enfin nous avons testé dans cette première étude de jeunes adolescents dans un contexte scolaire mixte, ce qui offre une bonne validité écologique à nos résultats. Soulignons également le fait que les adolescents sont rarement testés dans ce domaine, alors qu'il nous semble de première importance de prendre la mesure de l'impact du stéréotype à un âge où l'identité de sexe se construit. Ainsi, nos résultats contribuent avec d'autres à montrer que des régulations de la performance fondées sur les stéréotypes sont déjà à l'œuvre avant l'âge adulte (cf. Ambady et coll., 2001, pour la mise en évidence de ces régulations auprès de jeunes enfants).

Pour valider notre hypothèse de l'intervention d'un stéréotype interférent au cours du MRT, il était impératif de montrer que les sujets, et particulièrement les femmes, au minimum connaissent, voire même croient les stéréotypes de sexe dans le domaine visuo-spatial. Les résultats de l'étude 2 nous ont permis de répondre positivement sur ce point. Par rapport à l'ensemble de la littérature, nous avons été surpris de constater que si l'idée d'un tel stéréotype était généralement acceptée, elle n'avait jusqu'ici jamais été formellement démontrée avant notre étude (Halpern, 2000 ; Robert, 1990). Des activités comme les mathématiques, le bricolage, et certains sports qui nécessitent de bonnes compétences spatiales sont des activités effectivement perçues comme des domaines d'ex-

cellence masculine (Newcombe et coll., 1983 ; Signorella & Jamison, 1986), mais il nous a semblé moins évident que de telles croyances existaient concernant un test cognitif que l'on voit rarement hors des laboratoires. Notre étude 2 montre bien que les étudiants des deux sexes nourrissent des attentes compatibles avec l'idée d'une infériorité féminine en RM, avec une nuance néanmoins. Tout en s'auto-évaluant plus négativement que leurs homologues masculins, les jeunes femmes admettent semble-t-il plus difficilement l'infériorité de leur groupe de sexe. Ce profil de réponse confirme ce que nous observons dans les questionnaires post-tests des études 1 et 3, où nous posons les mêmes questions au sujet du MRT (après confrontation au test). On peut interpréter cette « résistance » comme un processus d'individuation où les femmes, tout en s'auto-évaluant moins positivement que les hommes, cherchent à protéger la réputation de leur groupe d'appartenance.

L'objectif de notre étude 3 était de confirmer les résultats de l'étude 1 chez des étudiantes de psychologie. Nous avons ainsi montré un effet positif de la falsification explicite du stéréotype, et aussi celui d'une intervention nouvelle forçant les sujets à se focaliser sur une comparaison en principe avantageuse avec les personnes âgées. Dans les deux cas ces interventions ont permis de prévenir la chute de performance observée par ailleurs (condition standard) au cours de la seconde section du test. Il reste qu'au niveau des scores globaux (performance sur la totalité du test) aucune différence n'était obtenue. On voit bien ici l'importance de ne pas négliger la dynamique temporelle à l'échelle du test puisque dans notre étude 3 l'interférence en question n'est visible qu'à ce niveau plus subtil.

Processus impliqués dans le stereotype threat

L'étude 4 apporte quant à elle quelques indications sur les mécanismes sous-jacents à la régulation objet de nos travaux. Tout d'abord cette étude a permis de valider le paradigme de falsification sur une version informatisée du test de RM (présentation par paire). Conformément à nos attentes, suite à la falsification du stéréotype, les étudiantes de

sciences répondaient plus rapidement que leurs homologues ou que les étudiantes de lettres dans toutes les autres conditions, sans que la justesse de leur réponse ne soit affectée. C'est à notre connaissance la première fois qu'une étude montre un effet positif de la falsification du stéréotype sur la vitesse de traitement de l'information. Dans notre étude, cet effet valait pour tous les types de problèmes, y compris lorsque l'amplitude des angles était égale à 0, donc lorsqu'aucune rotation mentale n'était requise. Cet autre résultat important permet de conclure à l'effet positif de la falsification en matière simplement décisionnelle, au moins lorsque l'activité cible consiste à décider de l'identité de deux objets. Cette observation heurte l'un des principes de base de la théorie du *stereotype threat*, qui fait de la difficulté du test une condition nécessaire à l'expression de l'interférence en question. Il semble au contraire que, du point de vue de la vitesse de traitement, l'effet puisse être observé même sur les tests *a priori* faciles. Il est possible que les effets d'interférence sur les tests faciles soient passés inaperçus en raison non seulement d'un désintérêt pour les mesures de vitesse, mais aussi parce que l'altération des temps de réponse n'est pas nécessairement toujours suffisante pour affecter *in fine* la performance à un niveau plus qualitatif.

Une innovation de notre travail repose sur une analyse originale (analyse distributionnelle) des temps de réponse recueillis dans cette étude 4. Conformément à nos attentes, cette analyse — inédite dans ce domaine — montre que l'effet positif de la falsification se traduit par une réduction des temps de réponse les plus longs, ce qui selon le modèle de Sanders (1983) démontre que l'effet observé est localisé au niveau des mécanismes énergétiques.

Notre analyse était guidée par une approche théorique qui, bien qu'ignorée par l'ensemble de la littérature sur le *stereotype threat*, nous a paru pertinente pour expliquer le déclin de la performance lié selon nous à l'intervention du stéréotype : le modèle du stress et de la performance de Sanders (1983). Ce modèle suppose l'existence de trois mécanismes de régulation du traitement de l'information : *arousal*, effort et activation. Ces

mécanismes sont associés aux ressources « énergétiques » (qualifiées comme telles parce que se référant à une ressource en quantité limitée) qui permettent une adaptation à la demande de la tâche et d'optimiser le traitement de l'information. L'effort, contrôlé par un mécanisme d'évaluation (notamment celle des réponses données au cours de la tâche), coordonne les deux autres mécanismes énergétiques afin d'éviter les déséquilibres. Selon nous, l'activation du stéréotype associée à la difficulté du test crée un enjeu pour les sujets ciblés par le stéréotype et qui ont les capacités de réussir, mais qui redoutent l'échec. Cela biaiserait en quelque sorte le mécanisme d'évaluation, avec pour conséquence un « sur-contrôle » de nature à augmenter l'effort investi dans la tâche mais dont l'issue serait, assez paradoxalement, une détérioration plutôt qu'une optimisation de la performance.

L'approche que nous défendons rejoint les hypothèses récentes qui situent l'interférence du stéréotype au niveau des processus d'auto-régulation, c'est-à-dire la capacité de contrôler et d'ajuster son propre comportement aux contraintes de l'activité cible et / ou de contexte de traitement. Ainsi Beilock et ses collaborateurs (2006, 2007 ; 2008) suggèrent que la mobilisation d'un stéréotype négatif peut en effet conduire les sujets à augmenter leurs efforts de contrôle dans l'activité cible, ce qui, dans le cas des tâches routinières ou fondées sur des automatismes, peut avoir pour effet d'entraver la performance. En détournant une partie de l'attention allouée par les sujets à une activité cible automatisée, Beilock et coll. parviennent à contrer l'effet délétère du stéréotype. Selon Ilkowska et coll. (2010), ces résultats suggèrent que l'augmentation de la charge cognitive consécutive à l'intervention du stéréotype diminue la capacité en mémoire de travail, et donc la capacité d'auto-régulation des sujets. Cette charge a sans doute une tonalité émotionnelle comme le laisse entrevoir les résultats récents obtenus dans ce domaine par IRMf (Krendl, Richeson, Kelley, & Heatherton, 2008 ; Wraga et coll., 2007). Dans ces travaux en effet, activer le stéréotype négatif donne lieu à des activations émotionnelles en particulier au niveau du cortex cingulaire antérieur ventral, région reconnue par ailleurs pour son implication dans le traitement de l'information émotionnelle (LeDoux, 2000). On sait par ailleurs que le cortex cingulaire antérieur est fortement impliqué dans les mécanismes de contrôle. Il

semble par exemple être le siège de la fameuse onde négative d'erreur dans les tâches impliquant une régulation des réponses dominantes (Dehaene, Posner, & Tucker, 1994 ; Veen & Carter, 2002). Les résultats d'Inzlicht et Kang (2010) vont dans le même sens à partir d'observation fondées sur la technique des potentiels évoqués (cf. Derks, Inzlicht, & Kang, 2008).

Notons que le modèle de Sanders (1983) lui-même n'est incompatible ni avec le modèle explicatif du *stereotype threat* présenté par Schmader et coll. (2008, tout en étant plus précis concernant les prédictions sur les temps de réponse par exemple), ni avec d'autres données suggérant l'implication de mécanismes tels que l'*arousal* (Ben-Zeev, Carrasquillo, et coll., 2005), ou l'augmentation de la charge cognitive, inférée à partir d'une réduction de la variabilité cardiaque (Croizet et coll., 2004). Il serait par conséquent utile à l'avenir de compléter notre approche dans le domaine des rotations mentales par des indicateurs neuro-physiologiques des processus impliqués.

Enjeux sociétaux

Nos travaux de thèse sur les différences de sexe dans le champ visuo-spatial présentent certains enjeux sociétaux.

De très nombreuses sociétés attribuent — et ont attribué —, des différences de capacités aux hommes et aux femmes sur la base des différences biologiques évidentes entre les sexes. Ces considérations, fondées ou non, justifient bien souvent l'assignation de certains rôles à l'un ou l'autre sexe. Aujourd'hui, la demande sociale est assez forte pour que les scientifiques déterminent non seulement la réalité des différences en question dans le domaine intellectuel, mais également l'origine biologique ou sociale de ces différences. Pour illustrer combien cette dernière question cristallise de tensions, rappelons la crise qui a suivi les déclarations — de son propre aveux maladroitement provocatrices — de Lawrence Summer, ex-président de l'Université Harvard. Lors d'une conférence en jan-

vier 2005, le président Summer déclara que s'il devait classer les facteurs expliquant la sur-représentation des hommes aux postes d'enseignement les plus distingués dans les disciplines scientifiques, il considérerait la meilleure « aptitude intrinsèque » des hommes pour les sciences avant de prendre en compte d'éventuelles discriminations³ (Ripley, 2005). Le tollé dans les médias et au sein de l'université fut suffisamment important pour aboutir un an après — bien que cela n'en soit certes pas la seule raison — à la démission du président Summer.

Si aujourd'hui tous acquiescent à l'idée que le débat inné / acquis est dépassé et que les positions doivent être plus nuancées (modèles bio-psycho-sociaux, cf. Casey, 1996b ; Halpern et coll., 2007), la déclaration du président Summer est néanmoins assez révélatrice. Elle semble relever de l'idée — somme toute assez répandue — qu'il existe une hiérarchie des facteurs explicatifs des différences de sexe dans le champ scientifique, dans laquelle les facteurs biologiques occupent une place prépondérante. La popularité de cette idée tient tant aux progrès récents en neurosciences, et notamment aux résultats obtenus grâce aux techniques d'imagerie cérébrale, qu'à la robustesse des effets de sexe observés dans les tâches concernées. Cette robustesse est telle, qu'intuitivement, on ne peut y reconnaître qu'une inscription biologique des différences de capacités. Par contraste, les études adossées classiquement à l'hypothèse de la socialisation peinent à déterminer quelle part de ces effets de sexe est expliquée par les facteurs biologiques et par les facteurs sociaux.

Plusieurs questions émergent de la même manière pour le paradigme du *stereotype threat* et du même coup pour nos propres travaux : les phénomènes observés dans nos recherches sont-ils importants ? Ces effets sont-ils antagonistes ou au contraire complémentaires avec les explications d'ordre biologique ? Il est clair que l'annulation des effets de sexe consécutive à la falsification du stéréotype remet en cause l'idée qu'une part de ces effets serait irréductible parce que due à des facteurs biologiques. En outre, ces derniers facteurs permettraient difficilement d'expliquer deux points marquants de nos résultats.

3. Le président Summer avait également déclaré que le facteur qui venait en premier était que les femmes refusaient de faire les sacrifices exigés par de tels postes.

D'une part nous montrons que les effets de cette falsification sont surtout sensibles chez les sujets de sexe féminin les plus performants en RM (cf. étude 1), et d'autre part que les effets de sexe observés résultent d'une érosion de la performance (prévenue par la falsification) au cours du test (cf. étude 1 et 3). Une telle malléabilité de la performance des sujets féminins, dans des proportions qui ne sont vraiment pas négligeables, n'est pas compatible avec une explication qui donnerait la primauté aux facteurs biologiques. Notons ici que la performance des hommes en RM est elle-même perméable à l'influence du stéréotype (*stereotype lift*), la sensibilité à ce dernier n'étant donc pas exclusive à un groupe de sexe en particulier.

Même en admettant que la taille des effets observés dans ce domaine soit assez modeste, on aurait tort de les négliger. Les effets de petite taille et *a fortiori* de taille modérée peuvent avoir des conséquences importantes au niveau de la population générale. Ainsi, Hyde (1981) établit pour différentes tailles d'effet de sexe les écarts théoriques entre les distributions de performance de chaque sexe et les sexe-ratios associés. Elle montre que dans le cas d'une taille d'effet de 5%, si l'on considère les individus au delà du 95ème percentile (les 5% des individus réalisant les meilleurs scores), le sexe-ratio est de deux hommes pour une femme. Or, théoriquement, ce sont précisément les femmes au plus haut niveau de la distribution statistique qui sont en principe les plus sensibles aux effets du stéréotype.

Une autre conséquence pratique très importante est la désaffection des femmes pour les filières scientifiques et techniques. En Europe, seules 36% de femmes obtiennent un doctorat dans ces domaines (European Commission, 2009). Là encore nous pouvons imaginer différents facteurs à l'origine de cette asymétrie, mais l'ensemble de la littérature sur l'interférence du stéréotype, à laquelle contribue notre travail de recherche, invite à reconsidérer le rôle joué spécifiquement par les stéréotypes de sexe. Bien sûr, l'effet délétère des stéréotypes n'est pas limité à la sphère académique. Par exemple, concernant un autre groupe social (des émigrés africains en Belgique), Klein et coll. (2007) ont montré

qu'un entretien d'embauche (simulé) constituait un contexte propice à la manifestation des effets négatifs d'un stéréotype sur les performances à un test d'intelligence. Les mécanismes impliquant les stéréotypes semblent donc intervenir à des étapes décisives pour la division du travail et sa hiérarchisation (Casini & Sanchez-Mazas, 2007 ; Klein & Pohl, 2007), ce qui n'est pas sans conséquence sur l'ensemble de notre organisation sociale.

Enfin, un mécanisme subtil peut éclairer encore la force des stéréotypes en matière de sélection des individus au sein de la hiérarchisation professionnelle. Il suffit de penser qu'au cours d'un cursus académique ou professionnel, un individu doit franchir une succession d'étapes de sélection, l'amenant à chaque fois à rejoindre un sous-groupe réunissant les plus performants. Or, nous l'avons vu, l'intervention d'un stéréotype est particulièrement dommageable aux femmes les plus performantes, ce qui réduit dramatiquement leurs chances de franchir chacune de ces étapes. Ces dernières se multipliant, les effets négatifs du stéréotype seront désastreux pour le nombre de femmes qui réussiront à franchir l'ensemble du processus de sélection, sans que, à une étape donnée, une détérioration de leurs performances ne soit nécessairement détectable (pour une analyse de ce phénomène au sujet de la promotion des femmes en milieu professionnel, cf. Martell, Lane, & Emrich, 1996 ; Valian, 1998).

Perspectives de recherche

Analyse distributionnelle. Concernant l'étude 4, nous avons réalisé des analyses distributionnelles en estimant la distribution des temps de réponse des sujets à l'aide d'une fonction ex-Gaussienne. Afin de réaliser un bon compromis entre la durée de la tâche et le nombre de TR mesurés, nous avons recueillis 40 TR par sujet et par condition. Si ce nombre est suffisant pour que l'algorithme converge dans la majorité des cas, il serait souhaitable de tripler au minimum le nombre d'observations afin d'augmenter dans de bonnes proportions la fiabilité de l'estimation des paramètres (Ratcliff, 1979 ; Cousineau et coll., 2004 ; Lacouture & Cousineau, 2008). Pour obtenir des distributions avec un nombre plus

important d'observations, nous pouvons envisager deux méthodes : (1) modifier notre version informatisée du MRT de manière à diminuer le nombre de modalités par sujet (essentiellement le nombre d'angles) et augmenter le nombre d'essais afin d'augmenter le nombre d'observations par modalité ; (2) mettre en œuvre la procédure du moyennage de Vincent (ou *vincentinisation*, Vincent, 1912 ; Ratcliff, 1979). Cette procédure inventée par Vincent en 1912 est très simple, il s'agit de regrouper les temps de réponse mesurés sur plusieurs sujets pour construire des distributions de TR de pseudo-sujets qui possèdent chacun un nombre d'observations plus important. Par exemple en regroupant trois sujets avec 40 observations chacun, nous créerions un pseudo-sujet associé à une distribution de TR constituée de 120 observations. Cette procédure est certes plus « coûteuse » en sujets, mais elle permet de contourner le problème du recueil d'un nombre d'observations suffisantes auprès d'un seul sujet, car cela nécessiterait des tâches trop longues pour être réalisées sérieusement. Dans tous les cas, si nous comptons développer cet axe de recherche, il faudra réaliser des analyses sur des distributions avec 100 observations au moins.

Stratégies. Nous envisageons par ailleurs de tester une hypothèse concernant une interaction entre l'activation du stéréotype et le type de stratégie mis en œuvre par les femmes. En effet, notre interprétation des résultats de l'étude 4 nous a amené à conclure que l'activation du stéréotype perturbait les mécanismes de contrôle d'allocation des ressources énergétiques, en raison probablement d'un « sur-contrôle » face à l'enjeu que constitue la situation de test pour les sujets menacés. Or l'hypothèse dominante dans la littérature sur les différences de sexe implique plutôt la manière dont les hommes et les femmes traitent l'information visuo-spatiale : les femmes adopteraient une stratégie analytique, moins efficace que la stratégie analogique privilégiée par les hommes, cette approche étant essentiellement étayée par les données de l'imagerie cérébrale (Thomsen et coll., 2000 ; Vogel et coll., 2003 ; Weiss et coll., 2003). Dans notre approche on ne peut exclure que l'adoption de cette stratégie analytique résulte au moins en partie de l'activation assez souvent impli-

cite du stéréotype dont les femmes font l'objet dans le domaine visuo-spatial. Une étude destinée à déterminer si la différence de sexe peut être expliquée par la mise en œuvre de stratégies différentes suite à l'activation du stéréotype nous semble donc nécessaire.

Pour pister ces stratégies, nous avons d'ores et déjà conçu un test de RM sur la base de celui créé pour l'étude 4. Nous envisageons dans ce nouveau test deux types de problèmes lorsque les objets sont différents. Dans un cas, les paires d'objets présentées diffèrent au niveau structurel (nombre et orientation des blocs constitutifs des objets), situation dans laquelle la stratégie analytique est efficace car elle consiste précisément à se focaliser sur la comparaison des éléments structurels des objets présentés. Dans l'autre cas, les deux objets sont classiquement l'image en miroir l'un de l'autre, situation assez ambiguë pour que la seule stratégie réellement efficace soit la stratégie analogique.

Stratégies et latéralisation. Plus généralement, de nombreuses études dans le domaine des différences de sexe prévoient d'intégrer des facteurs tant sociaux que biologiques de manière à comparer leurs effets respectifs ou leur éventuelle interaction (Casey, 1996b ; Halpern et coll., 2007). Pour le moment peu d'auteurs ont tenté une approche conjointe avec le paradigme du *stereotype threat*. À notre connaissance, seuls Wraga et coll. (2007, étude par IRMf) et Hausmann et coll. (2009, mesure du niveau de testostérone salivaire) l'ont fait. On notera pour notre part que certaines explications biologiques, en particulier l'hypothèse de latéralisation, sont typiquement examinées à l'aide d'observables comportementaux (p. ex., paradigme de la présentation en champ divisé). Il est par conséquent parfaitement envisageable de chercher à dissocier dans la différence de sexe au cœur de nos travaux ce qui relève d'une différence de latéralisation, d'une part, et ce qui relève de l'intervention d'un stéréotype menaçant, d'autre part. Notre hypothèse est que l'intervention du stéréotype négatif dont les femmes font l'objet dans le domaine visuo-spatial les conduit à mettre en œuvre une stratégie de type analytique au détriment de la stratégie analogique que l'on sait plus efficace.

Nous avons d'ores et déjà entrepris une étude préliminaire à ce sujet sur une centaine d'individus des deux sexes. Les résultats ne sont pour l'instant pas concluants, probablement en raison d'une difficulté trop grande des problèmes retenus dans cette étude. En effet le paradigme de présentation en champ divisé repose sur des contraintes fortes. Par exemple, les stimuli doivent être présentés très rapidement (une durée de l'ordre de 150 ms), c'est-à-dire avant même que les sujets ne déclenchent une saccade oculaire. Cette contrainte force par ailleurs à n'exposer les sujets qu'à des figures en deux dimensions, or nous l'avons dit antérieurement, l'effet de sexe n'est typiquement pas observé avec de telles figures. Dans notre expérience, cette contrainte « 2D » a été respectée tout en augmentant par conséquent la complexité des figures présentées aux sujets. Cet effort méthodologique s'est révélé cependant inefficace.

Contrôle cognitif et imagerie cérébrale. L'électro-encéphalographie est une technique qui nous apparaît prometteuse dans une perspective bio-psycho-sociale, notamment en rapport avec l'approche théorique que nous proposons dans cette thèse. Dans l'étude 4 nous avons montré la pertinence de la théorie du stress et de la performance de Sanders (1983) pour éclairer les processus sous-jacents à l'interférence du stéréotype. D'après nos résultats, il semblerait que ce que Sanders (1983) désigne comme les mécanismes énergétiques — que d'autres désignent sous le terme générique d'effort (Kahneman, 1973) ou de contrôle attentionnel (Posner & Boies, 1971) — soient en jeu. Dans le modèle de Sanders, l'effort, sous le contrôle d'un système d'évaluation, est un mécanisme dont la fonction est de coordonner les autres mécanismes énergétiques (*arousal* et activation) pour éviter les déséquilibres engendrés par des perturbations de différentes natures, ici l'intervention du contexte social sous la forme d'un stéréotype. Les résultats de notre étude 4 suggèrent que l'activation du stéréotype en effet perturbe ce mécanisme d'allocation des ressources en créant un phénomène de « sur-contrôle ».

Les techniques d'imagerie cérébrale telles que l'EEG ou l'IRMf nous permettraient de tester précisément cette hypothèse. En effet, l'EEG permet d'observer, lorsqu'un su-

jet commet une erreur, une onde de polarité négative, appelée *négativité d'erreur* ou *onde d'erreur* qui typiquement survient environ 70 ms après la production cette erreur (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991 ; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000 ; Gehring, Goss, Coles, Meyer, & Donchin, 1993). L'onde en question semble témoigner de l'activité du cortex cingulaire antérieur (CCA), structure impliquée dans les mécanismes d'évaluation et de détection de l'erreur (Dehaene et coll., 1994 ; Veen & Carter, 2002). Le sur-contrôle que nous supposons chez les femmes suite à l'activation du stéréotype devrait donc être observable via la mesure de cette onde d'erreur, par rapport à une condition où le stéréotype est falsifié. En effet des résultats montrent que chez des sujets particulièrement engagés dans le contrôle (patients atteints de troubles obsessionnel-compulsifs) l'onde de négativité d'erreur est plus ample que chez le sujet sain (Gehring, Himle, & Nisenson, 2000). Par ailleurs, une hyperactivité du CCA a été rapportée chez de tels patients (Dimond Fitzgerald et coll., 2005). Notre raisonnement est également étayé par un résultat très récent d'Inzlicht et Kang (2010) montrant en effet que l'activation d'un stéréotype négatif augmente l'amplitude des ondes d'erreurs. Cette observation isolée reste néanmoins à confirmer. La signification même de l'onde d'erreur reste à approfondir, comme le montrent des débats récents dans ce domaine (Burle, Roger, Allain, Vidal, & Hasbroucq, 2008 ; Roger, Bénar, Vidal, Hasbroucq, & Burle, 2010).

Conclusion

En somme, nos travaux de thèse confortent selon nous l'intérêt d'interroger le fondement social d'une différence de sexe encore souvent considérée par ailleurs comme quasi-naturelle. Loin de minimiser l'ampleur des travaux encore nécessaires pour valider notre thèse, il nous semble que les résultats dans ce cadre sont encourageants et fournissent d'ores et déjà des éléments importants (nous ne disons pas décisifs) pour l'avenir dans ce domaine de recherche. Non seulement la différence de sexe au cœur de nos travaux apparaît-elle malléable sous l'effet du contexte, mais ce contexte fait lui-même référence

à des connaissances sociales, donc hautement intégrées, de nature à interférer avec des processus d'assez bas-niveau et encore souvent appréhendés indépendamment de la dimension sociale du sujet. On voit bien ici tout l'intérêt d'appréhender la cognition dans son contexte social. À ce titre nos travaux complètent ceux inscrits dans le cadre d'une psychologie sociale de la cognition (Monteil & Huguet, 1999 ; Higgins, 2000).

Références

- Albaret, J.-M., & Aubert, E. (1996). Etalonnage 15-19 ans du test de rotation mentale de vanderberg. *Evolutions psychomotrices*, 8(34), 206-215.
- Alington, D. E., Leaf, R. C., & Monaghan, J. R. (1992). Effects of stimulus color, pattern, and practice on sex differences in mental rotations task performance. *Journal of Psychology : Interdisciplinary and Applied*, 126(5), 539-553.
- 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., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children : Effects of identity activation on quantitative performance. *Psychological Science*, 12(5).
- Aronson, J., Cohen, G. L., McColskey, W., & Regional Educational Laboratory Southeast. (2009). *Reducing stereotype threat in classrooms : A review of social-psychological intervention studies on improving the achievement of black students. issues & answers. rel 2009-076* (Rap. Tech.). Greensboro, NC: Regional Educational Laboratory Southeast.
- 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, 29-46.
- Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance by men and women in a virtual morris water task : A large and reliable sex difference (english). *Behavioural brain research*, 93(1-2), 185-190.
- Baddeley, A. D. (1986). *Working memory*. New York, NY US: Clarendon Press/Oxford University Press.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7(2), 85-97.
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science (Wiley-Blackwell)*, 17(5), 339-343.
- 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), 256-276.

-
- Bem, S. L. (1981). Gender schema theory : A cognitive account of sex typing. *Psychological Review*, 88(4), 354-364.
- Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents : their nature, effects, and possible causes (english). *Behavioral and brain sciences (Print)*, 11(2), 169-232.
- Ben-Zeev, T., Carrasquillo, C. M., Ching, A. M. L., Kliengklom, T. J., McDonald, K. L., Newhall, D. C., et coll. (2005). « math is hard ! » (barbietm, 1994) : Responses of threat vs. challenge-mediated arousal to stereotypes alleging intellectual inferiority. Dans M. Gallagher Ann & C. Kaufman James (Eds.), *Gender differences in mathematics : An integrative psychological approach*. Cambridge, MA: Cambridge University Press.
- Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, 41, 174-181.
- Blakemore, J. E. O., Berenbaum, S. A., & Liben, L. S. (2009). *Gender development*. New York, NY US: Psychology Press.
- Blascovich, J., Spencer, S. J., Quinn, D. M., & Steele, C. M. (2001). African americans and high blood pressure : The role of stereotype threat. *Psychological Science*, 12(3), 225-229.
- Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict : A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, 20(9), 1637-1655.
- Butler, T., Imperato-McGinley, J., Pan, H., Voyer, D., Cordero, J., Zhu, Y.-S., et coll. (2006). Sex differences in mental rotation : Top-down versus bottom-up processing. *NeuroImage*, 32, 445-456.
- Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat ? *Psychological Science*, 16(7), 572-578.
- Casey, M. B. (1996a). Gender, sex, and cognition : Considering the interrelationship between biological and environmental factors. *Learning and Individual Differences*, 8(1), 39-53.
- Casey, M. B. (1996b). Understanding individual differences in spatial ability within females : A nature/nurture interactionist framework. *Developmental Review*, 16, 241-260.
- Casey, M. B., Brabeck, M. M., & Nuttall, R. L. (1995). As the twig is bent : The biology and socialization of gender roles in women. *Brain and Cognition*, 27(2), 237-246.
- Casey, M. B., Nuttall, R. L., & Pezaris, E. (1999). Evidence in support of a model that predicts how biological and environmental factors interact to influence spatial skills. *Developmental Psychology*, 35(5), 1237-1247.
- Casey, M. B., Nuttall, R. L., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, 31(4), 697-705.
- Casini, A., & Sanchez-Mazas, M. (2007). Un regard à travers le « plafond de verre » : les mécanismes psychosociaux impliqués dans la sous-représentation des femmes au sommet de la hiérarchie professionnelle. Dans O. Klein & S. Pohl (Eds.), *Psychologie des stéréotypes et des préjugés* (p. 241-265). Loverval: Labor.
- Castelli, L., Latini Corazzini, L., & Carlo Geminiani, G. (2008). Spatial navigation in large-scale virtual environments : Gender differences in survey tasks (english). *Computers in human behavior*, 24(4), 1643-1667.

-
- Cattell, R. B. (1971). *Abilities : their structure, growth, and action*. Oxford England: Houghton Mifflin.
- Chalabaev, A., Sarrazin, P., Stone, J., & Cury, F. (2008). Do achievement goals mediate stereotype threat ? : An investigation on females' soccer performance. *Journal of Sport & Exercise Psychology*, 30(2), 143-158.
- Cherney, I. (2008). Mom, let me play more computer games : They improve my mental rotation skills. *Sex Roles*, 59(11/12), 776-786.
- Cherney, I., & London, K. (2006). Gender-linked differences in the toys, television shows, computer games, and outdoor activities of 5- to 13-year-old children. *Sex Roles*, 54(9-10), 717-726.
- Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *Journal of Neuroscience*, 29(40), 12574-12583.
- Clements-Stephens, A. M., Rimrodt, S. L., & Cutting, L. E. (2009). Developmental sex differences in basic visuospatial processing : Differences in strategy use? *Neuroscience Letters*, 449(3), 155-160.
- Cochran, K. F., & Wheatley, G. H. (1988). Ability and sex-related differences in cognitive strategies on spatial tasks. *The Journal of General Psychology*, 116(1), 43-55.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap : A social-psychological intervention. *Science*, 313(5791), 1307-1310.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (rev. ed.)*. Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98-101.
- Cohen, M. S., Kosslyn, S. M., Breiter, H. C., Digirolamo, G. J., Thomson, W. L., Anderson, A. K., et coll. (1996). Changes in cortical activity during mental rotation : A mapping study using functional magnetic resonance imaging. *Brain*, 119, 89-100.
- Collins, D. W., & Kimura, D. (1997). A large sex difference on a two-dimensional mental rotation task. *Behavioral Neuroscience*, 111(4), 845-849.
- Corballis, M. C. (1989). Laterality and human evolution. *Psychological Review*, 96(3), 492-505.
- Corballis, M. C. (1997). Mental rotation and the right hemisphere. *Brain and Language*, 57(1), 100-121.
- Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences*. Hove, UK: Psychology Press.
- Cousineau, D. (1999). *Analyses et modèles des temps de réponse, de l'apprentissage, et de l'automatisation de la recherche mnésique et visuelle*. Phd, Université de Montréal.
- Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood : Methods and package. *Behavior Research, Methods, Instruments & Computers*, 36, 277-290.
- Cousineau, D., & Larochelle, S. (1997). Pastis : A program for curve and distribution analyses. *Behavior Research Methods, Instruments, & Computers*, 29(4), 542-548.
- 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.
- Croizet, J.-C., Després, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive

-
- mental load. *Personality and Social Psychology Bulletin*, 30(6), 721-731.
- Croizet, J.-C., & Dutrévis, M. (2004). Socioeconomic status and intelligence : Why test scores do not equal merit. *Journal of Poverty*, 8(3), 91-107.
- Curtis, C. E. (2006). Prefrontal and parietal contributions to spatial working memory. *Neuroscience*, 139(1), 173-180.
- Dabbs, J. M., Chang, E.-L., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior*, 19, 89-98.
- Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, 38(6), 1639-1655.
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images : How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, 28(12), 1615-1628.
- Davranche, K., Audiffren, M., & Denjean, A. (2006). A distributional analysis of the effect of physical exercise on a choice reaction time task. *Journal of Sports Sciences*, 24(3), 323-329.
- Deaux, K. (1985). Sex and gender. *Annual Review of Psychology*, 36, 49-81.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, 5(5), 303-305.
- Delgado, A. R., & Prieto, G. (1996). Sex differences in visuospatial ability : Do performance factors play such an important role ? *Memory and Cognition*, 24(4), 504-510.
- Delignières, D. (1993). Anxiété et performance. Dans J.-P. Famose (Ed.), *Cognition et performance* (p. 235-254). Paris: INSEP.
- De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior*, 12(3), 351-361.
- De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology*, 163(3), 272-282.
- Derks, B., Inzlicht, M., & Kang, S. (2008). The neuroscience of stigma and stereotype threat. *Group Processes & Intergroup Relations*, 11(2), 163-181.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & Knippenberg, A. van. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, 36, 531-544.
- Dimond Fitzgerald, K., Welsh, R. C., Gehring, W. J., Abelson, J. L., Himle, J. A., Liberzon, I., et coll. (2005). Error-related hyperactivity of the anterior cingulate cortex in obsessive-compulsive disorder (english). *Biological psychiatry* (1969), 57(3), 287-294.
- Dror, I. E., Schmitz-Williams, I. C., & Smith, W. (2005). Older adults use mental representations that reduce cognitive load : Mental rotation utilizes holistic representations and processing. *Experimental Aging Research*, 31(4), 409-420.
- Désert, M., Croizet, J.-C., & Leyens, J.-P. (2002). La menace du stéréotype : Une interaction entre situation et identité. *L'Année Psychologique*, 102, 555-576.
- Duff, S. J., & Hampson, E. (2000). A beneficial effect of estrogen on working memory in postmenopausal women taking hormone replacement therapy. *Hormones and Behavior*, 38, 262-276.

-
- Eals, M., & Silverman, I. (1994). The hunter-gatherer theory of spatial sex differences : Proximate factors mediating the female advantage in recall of object arrays. *Ethology & Sociobiology*, 15(2), 95-105.
- Eme, P. E., & Marquer, J. (1998). Quantitative and qualitative individual differences in visual comparison processing : are strategies only analytic or holistic ? *Personality & Individual Differences*, 25, 459-475.
- European Commission. (2009). *She figures 2009 : Statistics and indicators on gender equality in science*. Luxembourg: Publications Office of the European Union.
- Fairweather, H., & Beaumont, J. G. (1982). Sex differences : Little reason for females to play midfield. Dans *Divided visual field studies of cerebral organisation* (p. 147-194). London: Academic.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks. *Electroencephalography And Clinical Neurophysiology*, 78, 447-455.
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). Erp components on reaction errors and their functional significance : a tutorial (english). *Biological psychology*, 51(2-3), 87-107.
- Feingold, A. (1992). Sex differences in variability in intellectual abilities : A new look at an old controversy. *Review of Educational Research*, 62(1), 61-84.
- Francis, M. A., & Irwin, R. J. (1997). Cerebral asymmetry and decision strategies in mental rotation : A psychophysical analysis. *European Journal of Cognitive Psychology*, 9(2), 225-240.
- French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Manual for kit of reference tests for cognitive factors (rev. ed.)*. Princeton, NJ: Educational Testing Service.
- Galea, L. A. M., Uban, K. A., Epp, J. R., Brummelte, S., Barha, C. K., Wilson, W. L., et coll. (2008). Endocrine regulation of cognition and neuroplasticity : Our pursuit to unveil the complex interaction between hormones, the brain, and behaviour. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 62(4), 247-260.
- Geary, D. C. (1995). Sexual selection and sexual differences in spatial cognition. *Learning and Individual Differences*, 7(4), 289-301.
- Geary, D. C. (1998). *Male, female : The evolution of human sex differences*. Washington, DC: American Psychological Association.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science (Wiley-Blackwell)*, 4(6), 385-390.
- Gehring, W. J., Himle, J. A., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, 11(1), 1-6.
- Geiser, C., Lehmann, W., & Eid, M. (2006). Separating "rotators" from "non-rotators" in the mental rotations test : A multigroup latent class analysis. *Multivariate Behavioral Research*, 41(3), 261-293.
- Gizewski, E. R., Krause, E., Wanke, I., Forsting, M., & Senf, W. (2006). Gender-specific cerebral activation during cognitive tasks using functional mri : comparison of women in mid-luteal phase and men. *Neuroradiology*, 48, 14-20.
- Glück, J., & Fitting, S. (2003). Spatial strategy selection : Interesting incremental information. *International Journal of Testing*, 3(3), 293-308.

-
- Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual-spatial ability : The role of performance factors. *Memory and Cognition*, 18(5), 546-550.
- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of latino women. *Personality and Social Psychology Bulletin*, 28(5), 659-670.
- Gray, J. A. (1971). *The psychology of fear and stress*. New York, NY England: Mcgraw-Hill Book Company.
- Guimond, S., Branscombe, N. R., Brunot, S., Buunk, A. P., Chatard, A., Désert, M., et coll. (2007). Culture, gender, and the self : Variations and impact of social comparison processes. *Journal of Personality and Social Psychology*, 92(6), 1118-1134.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, New Jersey: Laurence Erlbaum Associates, Inc.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8(1), 1-51.
- Halpern, D. F., & Tan, U. (2001). Stereotypes and steroids : Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cognition*, 45, 392-414.
- Halpern, D. F., & Wright, T. M. (1996). A process-oriented model of cognitive sex differences. *Learning and Individual Differences*, 8(1), 3-24.
- Hamilton, C. J. (1995). Beyond sex differences in visuo-spatial processing : the impact of gender trait possession. *British Journal of Psychology*, 86, 1-20.
- Harshman, R. A., Hampson, E., & Berenbaum, S. A. (1983). Individual differences in cognitive abilities and brain organization : I. sex and handedness differences in ability. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 37(1), 144-192.
- Hausmann, M., Schoofs, D., Rosenthal, H. E. S., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences—a psychobiosocial approach. *Psychoneuroendocrinology*, 34(3), 389-401.
- Hausmann, M., Slabbekoorn, D., Van Goozen, H. M., Cohen-Kettenis, P. T., & Güntürkün, O. (2000). Sex hormones affect spatial abilities during the menstrual cycle. *Behavioral Neuroscience*, 114(6), 1245-1250.
- Healy, S. D., Braham, S. R., & Braithwaite, V. A. (1996). Spatial working memory in rats : no differences between the sexes. *Proceedings of the Royal Society of London, B*, 266, 2303-2308.
- Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions : An exemple using the stroop task. *Psychological Bulletin*, 109(2), 340-347.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269(5220), 41-45.
- Heilman, M. E., Block, C. J., & Martell, R. F. (1995). Sex stereotypes : Do they influence perceptions of managers ? *Journal of Social Behavior & Personality*, 10(6), 237-252.
- Higgins, E. T. (2000). Social cognition : learning about what matters in the social world. *European Journal of Social Psychology*, 30(1), 3-39. (Accession Number : 419424; Authors : Higgins, E. Tory A1)
- Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, 99(3), 545-560.

-
- 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.
- Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance : A theory of the relationship between individual differences and information processing. *Psychological Review*, 91(2), 153-184.
- Hyde, J. S. (1981). How large are cognitive gender differences ? a meta-analysis using ω^2 and d . *American Psychologist*, 36(8), 892-901.
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability : A meta-analysis. *Psychological Bulletin*, 104(1), 53-59.
- Ilkowska, M., Engle, R. W., & Hoyle, R. H. (2010). Working memory capacity and self-regulation. Dans *Handbook of personality and self-regulation*. (p. 265-290). Oxford, UK: Wiley-Blackwell.
- Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover : How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology*, 99(3), 467-481.
- Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, 83(4), 518-527.
- Jahoda, G. (1980). Sex and ethnic differences on a spatial-perceptual task : Some hypotheses tested. *British Journal of Psychology*, 71(3), 425.
- James, T. W., & Kimura, D. (1997). Sex differences in remembering the locations of objects in an array : Location-shifts versus location-exchanges. *Evolution and Human Behavior*, 18(3), 155-163.
- Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion : Examining the influence of emotion regulation. *Journal of Experimental Psychology : General*, 137(4), 691-705.
- Johnson, B. W., McKenzie, K. J., & Hamm, J. P. (2002). Cerebral asymmetry for mental rotation : effects of response hand, handedness and gender. *Neuroreport : For Rapid Communication of Neuroscience Research*, 13(15), 1929-1932.
- Jordan, K., Schadow, J., Wüstenberg, T., Heinze, H.-J., & Jäncke, L. (2004). Different cortical activations for subjects using allocentric or egocentric strategies in a virtual navigation task. *Brain Imaging*, 15(1), 135-140.
- Jordan, K., Wüstenberg, T., Heinze, H.-J., Peters, M., & Jäncke, L. (2002). Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia*, 40, 2397-2408.
- Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems : Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92(2), 137-172.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kanamori, N., & Yagi, A. (2002). The difference between flipping strategy and spinning strategy in mental rotation. *Perception*, 31(12), 1459-1466.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., Engle, R. W., Miyake, A., Jarrold, C., et coll. (2007). Variation in working memory capacity as variation in executive attention and control. Dans *Variation in working memory*. (p. 21-46). New York, NY US: Oxford University Press.
- Keenan, P. A., Ezzat, W. H., Ginsburg, K., & Moore, G. J. (2001). Prefrontal cortex as the site of estrogen's effect on cognition. *Psychoneuroendocrinology*, 26, 577-590.

-
- 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., & 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.
- Keller, J., & Sekaquaptewa, D. (2008). Solo status and women's spatial test performance : the role of individuation tendencies. *European Journal of Social Psychology*, 38(6), 1044-1053.
- Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.
- Kimura, D. (2004). Human sex differences in cognition, fact, not predicament. *Sexualities, Evolution & Gender*, 6(1), 45-53.
- Klein, O., & Pohl, S. (2007). Les biais rencontrés lors du processus de sélection professionnelle : une analyse en termes de stéréotypes. Dans O. Klein & S. Pohl (Eds.), *Psychologie des stéréotypes et des préjugés*. (p. 197-239). Lovreval: Labor.
- Klein, O., Pohl, S., & Ndagijimana, C. (2007). The influence of intergroup comparisons on africans' intelligence test performance in a job selection context. *Journal of Psychology*, 141(5), 453-468.
- Kosslyn, S. M. (1981). The medium and the message in mental imagery : A theory. *Psychological Review*, 88(1), 46-66.
- 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.
- Kringelbach, M. L. (2005). The human orbitofrontal cortex : Linking reward to hedonic experience (english). *Nature reviews. Neuroscience (Print)*, 6(9), 691-702.
- Kuhn, D., Nash, S. C., & Brucken, L. (1978). Sex role concepts of two- and three-year-olds. *Child Development*, 49, 445-451.
- Kuhn, S. L., & Stiner, M. C. (2006). What's a mother to do ? : The division of labor among neandertals and modern humans in eurasia (english). *Current anthropology*, 47(6), 953-980.
- Lacouture, Y., & Cousineau, D. (2008). How to use matlab to fit the ex-gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*, 4(1), 35-45.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1), 155.
- Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of adhd children : a response time distributional approach. *Acta Psychologica*, 104(2), 167-190.
- Leyens, J.-P., Désert, M., Croizet, J.-C., & Darcis, C. (2000). Stereotype threat : Are lower status and history of stigmatization preconditions of stereotype threat ? *Personality and Social Psychology Bulletin*, 26(10), 1189-1199.
- Leyens, J.-P., Yzerbyt, V., & Schadron, G. (1994). *Stereotypes and social cognition*. Thousand Oaks, CA US: Sage Publications, Inc.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability : A meta-analysis. *Child Development*, 56, 1479-1498.

-
- LoPresti, M. L., Schon, K., Tricarico, M. D., Swisher, J. D., Celone, K. A., & Stern, C. E. (2008). Working memory for social cues recruits orbitofrontal cortex and amygdala : A functional magnetic resonance imaging study of delayed matching to sample for emotional expressions. *Journal of Neuroscience*, 28(14), 3718-3728.
- Loring-Meier, S., & Halpern, D. F. (1999). Sex differences in visuospatial working memory : Components of cognitive processing. *Psychonomic Bulletin and Review*, 6(3), 464-471.
- Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years : Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, 1(4), 316-345.
- Luce, R. D. (1986). *Response times : Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Maccoby, E. E. (1966). *Sex differences in intellectual functioning*. (E. E. Maccoby, Ed.). Stanford, CA: Stanford University Press.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- Martell, R. F., Lane, D. M., & Emrich, C. (1996). Male-female differences : A computer simulation. *American Psychologist*, 51(2), 157-158.
- 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, 236-243.
- Martinot, D., & Redersdorff, S. (2003). Impact of comparisons with out-group members on women's self-esteem : Role of the stereotypical connotation of the performance context. *International Journal of Psychology*, 38(6), 348-358.
- Masters, M. S. (1998). The gender difference on the mental rotations test is not due to performance factors. *Memory and Cognition*, 26(3), 444-448.
- Mayer, D. M., & Hanges, P. J. (2003). Understanding the stereotype threat effect with « culture-free » tests : An examination of its mediators and measurement. *Human Performance*, 16(3), 207-230.
- McCarthy, D. (1954). Language development in children. Dans L. Carmichael (Ed.), *Manual of child psychology*. (2d ed., p. 476-581). Hoboken, NJ US: John Wiley & Sons Inc.
- McGee, M. G. (1978). Effect of two problem solving strategies on mental rotation test scores. *Journal of Psychology*, 100(1), 83-86.
- McGlone, J. (1978). Sex differences in functional brain asymmetry. *Cortex; A Journal Devoted To The Study Of The Nervous System And Behavior*, 14(1), 122-128.
- McGlone, J. (1980). Sex differences in human brain asymmetry : A critical survey. *Behavioral and Brain Sciences*, 3(2), 215-263.
- McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, 27(5), 486-493.
- McKeever, W. F. (1995). Hormone and hemisphericity hypotheses regarding cognitive sex differences : Possible future explanatory power, but current empirical chaos. *Learning and Individual Differences*, 7(4), 323-340.
- McMahan, I. D. (1982). Expectancy of success on sex-linked tasks. *Sex Roles : A Journal of Research*, 8(9), 949-58.
- Meehan, A. M. (1984). A meta-analysis of sex differences in formal operational thought. *Child Development*, 55(3), 1110-1124.

-
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? a latent-variable analysis. *Journal of Experimental Psychology : General*, 130(4), 621-640.
- Moè, A., & Pazzaglia, F. (2006). Following the instructions! : Effects of gender beliefs in mental rotation. *Learning and Individual Differences*, 16(4), 369-377.
- Monteil, J.-M., & Huguet, P. (1999). *Social context and cognitive performance : Towards a social psychology of cognition*. Hove England: Psychology Press/Taylor & Francis (UK).
- Moody, M. S. (1998). *Problem-solving strategies used on the mental rotations test : Their relationship to test instructions, scores, handedness, and college major*. Thèse non publiée.
- Murray, J. E. (1997). Flipping and spinning : Spatial transformation procedures in the identification of rotated natural objects. *Memory and Cognition*, 25(1), 96-105.
- Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics : Attitudes and stereotype threat susceptibility in italian children. *Developmental Psychology*, 43(3), 747-759.
- Newcombe, N., Bandura, M. M., & Taylor, D. G. (1983). Sex differences in spatial ability and spatial activities. *Sex Roles*, 9(3), 377-386.
- Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? a meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93(6), 1314-1334.
- 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.
- Osborne, J. W. (2001). Testing stereotype threat : Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, 26(3), 291-310.
- Owen, A. M., Herrod, N. J., Menon, D. K., Clark, J. C., Downey, S. P. M. J., Carpenter, T. A., et coll. (1999). Redefining the functional organization of working memory processes within human lateral prefrontal cortex. *European Journal of Neuroscience*, 11(2), 567-574.
- Owen, L. (2005). *Distorting the past. gender and the division of labor in the european palaeolithic*. Tübingen: Kerns verlag.
- Peters, M. (2005). Sex differences and the factor of time in solving vanderberg and kuse mental rotation problems. *Brain and Cognition*, 57(2), 176-184.
- Peters, M., & Battista, C. (2008). Applications of mental rotation figures of the shepard and metzler type and description of a mental rotation stimulus library. *Brain & Cognition*, 66(3), 260-264.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn vanderberg and kuse mental rotations test : Different versions and factors that affect performance. *Brain and Cognition*, 28(1), 39-58.
- Peters, M., Lehmann, W., Takahira, S., Takeuchi, Y., & Jordan, K. (2006). Mental rotation test performance in four cross-cultural samples (n = 3367) : Overall sex differences and the role of academic program in performance. *Cortex (Science Direct)*, 42(7), 1005-1014.
- Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York, NY: W. W. Norton.

-
- Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78(5), 391-408.
- Pribram, K. H., & McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review*, 82(2), 116-149.
- Prinzel, L. J. I., & Freeman, F. G. (1995). Sex differences in visuo-spatial ability : Task difficulty, speed-accuracy tradeoff, and other performance factors. *Canadian Journal of Experimental Psychology*, 49(4), 530-539.
- Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality & Individual Differences*, 40(3), 609-619.
- Quaiser-Pohl, C., & Lehmann, W. (2002). Girls' spatial abilities : Charting the contributions of experiences and attitudes in different academic groups. *British Journal of Educational Technology*, 72, 245-260.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86(3), 446-461.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510-532.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for raven's progressive matrices and vocabulary scales*. San Antonio, TX: Harcourt Assessment.
- Redersdorff, S., & Martinot, D. (2003). Impact des comparaisons ascendantes et descendantes sur l'estime de soi : Importance de l'identité mise en jeu (french). *L'Année Psychologique*, 103(3), 411-443.
- Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J.-M., & Huguet, P. (in press). Individual differences in working memory moderate stereotype threat effects. *Psychological Science*..
- Richardson, J. T. E., & Vecchi, T. (2002). A jigsaw-puzzle imagery task for assessing active visuospatial processes in old and young people. *Behaviour Research Methods, Instruments, & Computers*, 34(1), 69-82.
- Rilea, S. L., Roskos-Ewoldsen, B., & Boles, D. (2004). Sex differences in spatial ability : A lateralization of function approach. *Brain and Cognition*, 56, 332-343.
- Ripley, A. (2005). Who says a woman can't be einstein? *Time*, 51-60.
- Robert, M. (1990). Sex-typing of the water-level task : There is more than meets the eye. *International Journal of Psychology*, 25(4), 475-490.
- Robert, M., & Savoie, N. (2006). Are there gender differences in verbal and visuospatial working-memory resources? *European Journal of Cognitive Psychology*, 18(3), 378-397.
- Roberts, J. E., & Bell, M. A. (2003). Two- and three-dimensional mental rotation tasks lead to different parietal laterality for men and women. *International journal of psychophysiology*, 50, 235-246.
- Roger, C., Bénar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral cingulate zone and correct response monitoring : Ica and source localization evidences for the unicity of correct- and error-negativities. *NeuroImage*, 51(1), 391-403.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom : Teacher expectation and pupils' intellectual development*. New York, NY US: Holt, Rinehart & Winston.
- Routtenberg, A. (1968). The two-arousal hypothesis : Reticular formation and limbic system. *Psychological Review*, 75(1), 51-80.

-
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53(1), 61-97.
- Scali, R. M., & Brownlow, S. (2000). Gender differences in spatial task performance as a function of speed or accuracy orientation. *Sex Roles*, 43(5-6), 359-376.
- Schimmel, J., Arndt, J., Banko, M., Katherine, & Cook, A. (2004). Not all self-affirmations were created equal : The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition*, 22(1), 75-99.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85(3), 440-452.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115(2), 336-356.
- Schmitzer-Torbert, N. (2007). Place and response learning in human virtual navigation : Behavioral measures and gender differences. *Behavioral Neuroscience*, 121(2), 277-290.
- Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies : Their effects on women's performance. *Journal of Experimental Social Psychology*, 39(1), 68.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing : An individual differences approach. *Journal of Experimental Psychology : General*, 125(1), 4-27.
- Shapiro, J., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats : Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11(2), 107-130.
- Sharps, M. J., Price, J. L., & Williams, J. K. (1994). Spatial cognition and gender : Instructional and stimulus influences on mental image rotation performance. *Psychology of Women Quarterly*, 18, 413-425.
- Sharps, M. J., Welton, A. L., & Price, J. L. (1993). Gender and task in the determination of spatial cognitive performance. *Psychology of Women Quarterly*, 17, 71-83.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171, 701-704.
- Shih, M., Bonam, C., Sanchez, D., & Peck, C. (2007). The social construction of race : Biracial identity and vulnerability to stereotypes. *Cultural Diversity and Ethnic Minority Psychology*, 13(2), 125-133.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility : Identity salience and shifts in quantitative performance. *Psychological Science*, 10(1), 80-83.
- Shiina, K., Saito, T., & Suzuki, K. (1997). Analysis of problem solving process of a mental rotations test - performance in shepard-metzler tasks. *Journal of Geometry and Graphics*, 1(2), 185-193.
- Siegel-Hinson, R. I., & McKeever, W. F. (2002). Hemispheric specialisation, spatial activity experience, and sex differences on tests mental rotation ability. *Laterality*, 7(1), 59-74.
- Signorella, M. L., & Jamison, W. (1986). Masculinity, femininity, androgyny, and cognitive performance : A meta-analysis. *Psychological Bulletin*, 100(2), 207-228.
- Signorella, M. L., Jamison, W., & Krupa, M. H. (1989). Predicting spatial performance from gender stereotyping in activity preferences and in self-concept. *Developmental*

-
- Psychology*, 25(1), 89-95.
- Silverman, I., Eals, M., Barkow, J. H., Cosmides, L., & Tooby, J. (1992). Sex differences in spatial abilities : Evolutionary theory and data. Dans *The adapted mind : Evolutionary psychology and the generation of culture*. (p. 533-549). New York, NY US: Oxford University Press.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Springer, S. P., & Deutsch, G. (1998). *Left brain, right brain : Perspectives from cognitive neuroscience (5th ed.)*. New York, NY US: W H Freeman/Times Books/ Henry Holt & Co.
- Stafford, R. E. (1961). *Identical blocks, form aa*. University Park: Pennsylvania State University, Office of Student Affairs.
- Steele, C. M. (1997). A threat in the air : How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.
- Sternberg, S. (1969). The discovery of processing stages. *Acta Psychologica*, 30, 276-315.
- Stone, J. (2002). Battling doubt by avoiding practice : The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin*, 28(12), 1667-1678.
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, 77(6), 1213-1227.
- Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, 34(4), 665-693.
- Stricker, L. J., & Ward, W. C. (2008). Stereotype threat in applied settings re-examined : A reply. *Journal of Applied Social Psychology*, 38(6), 1656-1663.
- Stumpf, H. (1993). Performance factors and gender-related differences in spatial ability : Another assessment. *Memory and Cognition*, 21(6), 828-836.
- Tagaris, G. A., Kim, S.-G., Strupp, J. P., Andersen, P., Ugurbil, K., & Georgopoulos, A. P. (1997). Mental rotation studied by functional magnetic resonance imaging at high field (4 tesla) : Performance and cortical activation. *Journal of Cognitive Neurosciences*, 9(4), 419-433.
- Testart, A. (1986). *Essai sur les fondements de la division sexuelle du travail chez les chasseurs-cueilleurs*. Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Thomsen, T., Hugdahl, K., Ersland, L., Barndon, R., Lundervold, A., Smievoll, A. I., et coll. (2000). Functional magnetic resonance imaging (fmri) study of sex differences in a mental rotation task. *Medical Science Monitor*, 6(6), 1186-1196.
- Thurstone, L. L., & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monographs*, 2.
- Tomasino, B., & Rumiati, R. I. (2004). Effects of strategies on mental rotation and hemispheric lateralization : Neuropsychological evidence. *Journal of Cognitive Neurosciences*, 16(5), 878-888.

-
- Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. Dans W. G. Kinzey (Ed.), *The evolution of human behavior : Primate models* (p. 183-237). Albany, NY: SUNY Press.
- Unterrainer, J., Wranek, U., Staffen, W., Gruber, T., & Ladurner, G. (2000). Lateralized cognitive visuospatial processing : Is it primarily gender-related or due to quality of performance ? *Neuropsychobiology*, 41, 95-101.
- Valian, V. (1998). *Why so slow ? the advancement of women*. Cambridge, MA US: The MIT Press.
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-604.
- Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin and Review*, 7(3), 424-465.
- Vasta, R., Belongia, C., & Ribble, C. (1994). Investigating the orientation effect on the water-level task : Who ? when ? and why ? *Developmental Psychology*, 30(6), 893-904.
- Veen, V. van, & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, 14(4), 593-602.
- Vincent, S. B. (1912). The function of the viborissae in the behavior of the white rat. *Behavioral Monographs*, 1(5).
- Vogel, J. J., Bowers, C. A., & Vogel, D. S. (2003). Cerebral lateralization of spatial abilities : A meta-analysis. *Brain and Cognition*, 52, 197-204.
- Voyer, D. (1996). On the magnitude of laterality effects and sex differences in functional lateralities. *Laterality*, 1(1), 51-83.
- Voyer, D. (1997). Scoring procedure, performance factors, and magnitude of sex differences in spatial performance. *American Journal of Psychology*, 110(2), p259 - 276.
- Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory : A meta-analysis. *Psychonomic Bulletin & Review*, 14(1), 23-38.
- Voyer, D., Rodgers, M. A., & McCormick, P. A. (2004). Timing conditions and the magnitude of gender differences on the mental rotations test. *Memory and Cognition*, 32(1), 72-82.
- Voyer, D., & Saunders, K. A. (2004). Gender differences on the mental rotations test : a factor analysis. *Acta Psychologica*, 117, 79-94.
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities : A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117(2), 250-270.
- Walter, K. D., Roberts, A. E., & Brownlow, S. (2000). Spatial perception and mental rotation produce gender differences in cerebral hemovelocity : A ted study. *Journal of Psychophysiology*, 14, 37-45.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39(5), 456.
- Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability : A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, 99(2), 397-420.
- Weiss, E., Siedentopf, C. M., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C., et coll. (2003). Sex differences in brain activation pattern during a visuospatial cognitive task : a functional magnetic resonance imaging study in healthy volunteers.

-
- Neuroscience Letters*, 344, 169-172.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior : A review of possible mechanisms. *Psychological Bulletin*, 127(6), 797-826.
- Witt, S. D. (2000). The influence of peers on children's socialization to gender roles. *Early Child Development and Care*, 162, 1-7.
- Witt, S. D., & Burriss, K. G. (2000). The influence of television on children's gender role socialization. *Childhood Education*, 76(5), 322.
- Wraga, M., Duncan, L., Jacobs, E. C., Helt, M., & Church, J. A. (2006). Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychonomic Bulletin and Review*, 13(5), 813-819.
- Wraga, M., Helt, M., Jacobs, E. C., & Sullivan, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social Cognitive and Affective Neuroscience*, 2(1), 12-19.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation fo strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Index des auteurs

- Aarts, H., 184, 201
Abelson, J. L., 201
Agnoli, F., 80, 207
Albaret, J.-M., 84, 198
Alington, D. E., 32, 62, 118, 198
Allain, S., 195, 199
Ambady, N., 56, 59, 80, 184, 198, 209
Andersen, P., 210
Anderson, A. K., 200
Apfel, N., 59, 200
Arndt, J., 59, 209
Aronson, J., 11, 13, 14, 54–61, 63, 66, 70–72, 75, 78, 80, 82, 117, 144, 145, 181, 198, 206, 210
Astur, R. S., 33, 198
Aubert, E., 84, 198
Audiffren, M., 153, 201

Baddeley, A. D., 28, 198
Bandura, M. M., 49, 207
Banko, M., Katherine, 59, 209
Bargh, J. A., 184, 201
Barha, C. K., 202
Barkow, J. H., 29, 210
Barndon, R., 210
Battista, C., 154, 158, 178, 207
Beaumont, J. G., 41, 202
Beilock, S. L., 60, 146, 187, 198
Bell, M. A., 42, 208
Belongia, C., 120, 211
Bem, S. L., 47, 199
Benbow, C. P., 20, 47, 199, 203, 206, 211
Ben-Zeev, T., 56, 58, 59, 146, 150, 188, 199
Berenbaum, S. A., 16, 29, 199, 203
Blakemore, J. E. O., 16, 46, 62, 118, 199
Blanke, L., 195, 202
Blanton, H., 55, 203
Blascovich, J., 145, 199
Block, C. J., 127, 203

Bénar, C. G., 195, 208
Boies, S. J., 148, 194, 208
Boles, D., 41, 208
Bonam, C., 59, 209
Bossaerts, P., 12, 200
Bowers, C. A., 42, 211
Brabeck, M. M., 37, 199
Braham, S. R., 39, 203
Braithwaite, V. A., 39, 203
Brake, B., 29, 211
Branscombe, N. R., 203
Breiter, H. C., 200
Brown, J., 198
Brown, S., 165, 200
Brownlow, S., 62, 73, 209, 211
Brucken, L., 46, 205
Brummelte, S., 202
Brunot, S., 203
Bryden, M. P., 11, 211
Burle, B., 195, 199, 208
Burris, K. G., 47, 212
Butler, T., 32, 33, 199
Buunk, A. P., 203

Cadinu, M., 16, 56, 60, 89, 111, 146, 183, 199
Camarano, D. M., 50, 201
Carlo Geminiani, G., 33, 199
Carpenter, P. A., 31, 204
Carpenter, T. A., 207
Carr, T. H., 147, 198
Carrasquillo, C. M., 56, 58, 188, 199
Carter, C. S., 188, 195, 211
Casey, M. B., 20, 37, 62, 114, 118, 154, 189, 193, 199
Casini, A., 191, 199
Castelli, L., 33, 199
Cattell, R. B., 24, 200
Celone, K. A., 206
Chalabaev, A., 56, 200

-
- Chang, E.-L., 33, 201
Chatard, A., 203
Cherney, I., 46, 50, 200
Ching, A. M. L., 199
Christ, S., 195, 202
Christopoulos, G. I., 12, 43, 45, 200
Church, J. A., 11, 212
Claire, T., 56, 200
Clark, J. C., 207
Clements-Stephens, A. M., 32, 33, 200
Cochran, K. F., 31, 200
Cohen, G. L., 55, 59, 74, 198, 200, 211
Cohen, J., 22, 95, 200
Cohen, M. S., 30, 200
Cohen-Kettenis, P. T., 12, 203
Coles, M. G. H., 195, 202
Collins, D. W., 27, 62, 200
Conway, A. R. A., 204
Cook, A., 59, 209
Corballis, M. C., 40, 41, 200
Cordero, J., 199
Cornoldi, C., 29, 62, 200
Cosmides, L., 29, 210
Court, J. H., 120, 208
Cousineau, D., 165, 167, 191, 200, 205
Crandall, C. S., 110, 146, 201, 207
Croizet, J.-C., 55, 56, 60, 61, 145, 188, 200, 201, 205
Curtis, C. E., 30, 201
Cury, F., 56, 200
Cutting, L. E., 32, 200
- Dabbs, J. M., 33, 201
Danaher, K., 110, 201
Darcis, C., 56, 205
Darley, J. M., 56, 210
Dauenheimer, D., 60, 205
Davies, P. G., 127, 146, 201
Davranche, K., 153, 201
Deaux, K., 16, 201
Dehaene, S., 188, 195, 201
Deisenhammer, E. A., 211
Delgado, A. R., 36, 201
Delignières, D., 144, 145, 201
De Lisi, R., 50, 201
Denjean, A., 153, 201
Derks, B., 188, 201
Després, G., 200
Deutsch, G., 40, 210
DeVore, I., 39, 211
- Digirolamo, G. J., 200
Dijksterhuis, A., 184, 201
Dimond Fitzgerald, K., 195, 201
Dodson, J. D., 150, 212
Dolan, R. J., 12, 200
Donchin, E., 195, 202
Douglas, V. I., 153, 205
Downey, S. P. M. J., 207
Dror, I. E., 34, 201
Désert, M., 55, 56, 109, 184, 201, 203, 205
Duff, S. J., 40, 201
Duncan, L., 11, 212
Dutrévis, M., 56, 201
- Eals, M., 29, 38, 202, 210
Eid, M., 33, 202
Ekstrom, R. B., 27, 202
Eme, P. E., 31, 202
Emrich, C., 191, 206
Engle, R. W., 61, 204
Epp, J. R., 202
Ersland, L., 210
European Commission, 190, 202
Ezzat, W. H., 40, 204
- Fairweather, H., 41, 202
Falkenstein, M., 195, 202
Fein, S., 59, 146, 150, 199
Feingold, A., 81, 202
Fitting, S., 31, 202
Forbes, C., 17, 209
Forsting, M., 73, 202
Francis, M. A., 31, 202
Freeman, F. G., 35, 208
French, J. W., 27, 202
Friedman, N. P., 30, 207
- Galea, L. A. M., 40, 202
Garcia, J., 59, 200
Gauzins, M.-E., 200
Geary, D. C., 12, 29, 38, 202, 203
Gehring, W. J., 195, 201, 202
Geiser, C., 33, 50, 202, 208
Georgopoulos, A. P., 210
Gerhardstein, R., 127, 201
Gernsbacher, M. A., 203
Gimmig, D., 208
Ginsburg, K., 40, 204
Gizewski, E. R., 73, 202
Glück, J., 31, 202
Güntürkün, O., 12, 203

-
- Goldstein, D., 35, 36, 203
Gonzales, P. M., 55, 203
Good, C., 198
Goss, B., 195, 202
Gray, J. A., 148, 203
Greenberg, J., 11, 206
Gruber, T., 44, 211
Guimond, S., 48, 203
Gur, R. C., 203
- Haldane, D., 35, 203
Halpern, D. F., 11, 12, 16, 19, 21, 28–31, 37, 46, 48, 53, 54, 62, 72, 81, 110, 114, 118, 182, 184, 189, 193, 203, 206
Hambrick, D. Z., 204
Hamilton, C. J., 49, 53, 62, 118, 203
Hamm, J. P., 42, 204
Hampson, E., 29, 40, 201, 203
Hanges, P. J., 60, 206
Harshman, R. A., 29, 203
Hasbroucq, T., 195, 199, 208
Hausmann, M., 11, 12, 14, 39, 74, 75, 78, 80, 193, 203
Healy, S. D., 39, 203
Heathcote, A., 153, 165, 200, 203
Heatherton, T. F., 187, 205
Hedges, L. V., 21, 81, 91, 182, 203
Hegarty, M., 30, 207
Heilman, M. E., 127, 203
Heinze, H.-J., 12, 33, 204
Helt, M., 11, 212
Herrod, N. J., 207
Higgins, E. T., 196, 203
Himle, J. A., 195, 201, 202
Hofer, A., 211
Hohnsbein, J., 195, 202
Hoormann, J., 195, 202
Hoptman, M. J., 211
Hoyle, R. H., 61, 204
Hugdahl, K., 210
Huguet, P., 56, 58, 63, 80, 113, 118, 196, 200, 203, 204, 207, 208
Humphreys, M. S., 150, 204
Hyde, J. S., 22, 23, 182, 190, 203, 204
- Ilkowska, M., 61, 187, 204
Imperato-McGinley, J., 29, 199, 211
Inhelder, B., 120, 207
Inzlicht, M., 59, 146, 150, 188, 195, 199, 201, 204
Irwin, R. J., 31, 202
- Jacklin, C. N., 12, 19, 21–23, 54, 182, 206
Jackson, M., 207
Jacobs, E. C., 11, 212
Jacobs, J. E., 47, 204
Jacobson, L., 54, 208
Jahoda, G., 48, 204
James, T. W., 29, 204
Jamison, W., 48, 49, 53, 126, 185, 209
Jarrold, C., 204
Jellison, W. A., 146, 198
Jäncke, L., 12, 33, 204
Johns, M., 11, 17, 59, 60, 72, 146, 204, 206, 209
Johnson, B. W., 42, 204
Jordan, K., 11, 12, 33, 44, 48, 73, 203, 204, 207
Just, M. A., 31, 204
- Kahneman, D., 148, 194, 204
Kanamori, N., 31, 156, 204
Kane, M. J., 61, 204
Kang, S., 188, 201
Kang, S. K., 188, 195, 204
Keenan, P. A., 40, 204
Keller, J., 59, 60, 63, 205
Kelley, W. M., 187, 205
Keough, K., 198
Kiesner, J., 16, 199
Kim, A., 80, 198
Kim, S.-G., 210
Kimura, D., 12, 19, 27, 29, 38, 62, 200, 204, 205
King Elbaz, Z., 153, 205
Klein, O., 55, 190, 191, 205
Kliengklom, T. J., 199
Knippenberg, A. van, 184, 201
Kosslyn, S. M., 31, 200, 205
Krause, E., 73, 202
Kremser, C., 211
Krendl, A. C., 187, 205
Kringelbach, M. L., 12, 45, 205
Krupa, M. H., 49, 209
Kuhn, D., 46, 205
Kuhn, S. L., 39, 205
Kuse, A. R., 26–28, 33, 64, 65, 67, 76, 83, 84, 86, 120, 153, 154, 182, 211
- Lacouture, Y., 165, 191, 205
Ladurner, G., 44, 211
Laeng, B., 207
Lane, D. M., 191, 206

-
- Larochelle, S., 167, 200
Latham, K., 207
Latini Corazzini, L., 33, 199
Leaf, R. C., 32, 198
LeDoux, J. E., 187, 205
Lehmann, W., 33, 48–50, 62, 202, 207, 208
Leth-Steensen, C., 153, 205
Leyens, J.-P., 53, 55–57, 200, 201, 205
Liben, L. S., 16, 199
Liberzon, I., 201
Linn, M. C., 11, 12, 21–27, 182, 204, 205
London, K., 46, 200
LoPresti, M. L., 12, 45, 206
Loring-Meier, S., 30, 31, 206
Lubinski, D., 20, 206, 211
Luce, R. D., 165, 206
Lundervold, A., 210
Lustina, M. J., 198
Lynch, C. I., 56, 210

Maass, A., 16, 199
Maccoby, E. E., 12, 19, 21–23, 54, 182, 206
Marquer, J., 31, 202
Martell, R. F., 127, 191, 203, 206
Martens, A., 11, 14, 59, 67, 69, 70, 74, 75, 78, 79, 183, 206
Martinot, D., 129, 206, 208
Master, A., 59, 200
Masters, M. S., 36, 206
Mayer, D. M., 60, 206
McCarthy, D., 21, 206
McColskey, W., 59, 198
McConnell, A. R., 60, 146, 198
McCormick, P. A., 36, 211
McDonald, K. L., 199
McGee, M. G., 31, 206
McGlone, J., 40, 41, 206
McGlone, M. S., 11, 14, 70–72, 75, 78, 80, 206
McGuinness, D., 148, 208
McKeever, W. F., 39, 41, 42, 49, 206, 209
McKenzie, K. J., 42, 204
McMahan, I. D., 127, 206
Meehan, A. M., 29, 206
Menon, D. K., 207
Metzler, J., 27, 65, 68, 153, 154, 156, 178, 209
Mewhort, D. J., 153, 203
Meyer, D. E., 195, 202
Milun, R., 33, 201

Mitchell, C., 35, 203
Mitchell, J. P., 59, 198
Miyake, A., 28, 30, 204, 207, 209
Moè, A., 11, 14, 69–72, 75, 78, 79, 183, 207
Monaghan, J. R., 32, 198
Monteil, J.-M., 196, 207, 208
Moody, M. S., 32, 207
Moore, G. J., 40, 204
Méot, A., 200
Murray, J. E., 156, 207
Muzzatti, B., 80, 207

Nash, S. C., 46, 205
Ndagijimana, C., 55, 205
Neuberg, S. L., 55, 209
Newcombe, N., 49, 50, 118, 185, 207
Newhall, D. C., 199
Nguyen, H.-H. D., 109, 110, 207
Nisenson, L. G., 195, 202
Nowell, A., 21, 81, 91, 182, 203
Nuttall, R. L., 20, 37, 62, 114, 154, 199

O'Brien, L. T., 146, 207
Ortiz, M. L., 33, 198
Osborne, J. W., 60, 145, 207
Owen, A. M., 29, 207
Owen, L., 39, 207
Owen-Smith, A., 59, 198

Paik, S. K., 59, 198
Pan, H., 199
Pazzaglia, F., 11, 14, 69–72, 75, 78, 79, 183, 207
Peck, C., 59, 209
Peters, M., 12, 32, 36, 48–50, 67, 154, 158, 178, 183, 204, 207
Petersen, A. C., 11, 12, 21, 23–27, 182, 205
Petty, R. E., 55, 60, 109, 184, 212
Pezaris, E., 20, 114, 199
Piaget, J., 120, 207
Pittinsky, T. L., 56, 80, 198, 209
Pohl, S., 55, 191, 205
Popiel, S. J., 153, 203
Posner, M. I., 148, 188, 194, 201, 208
Postma, A., 29, 211
Pribram, K. H., 148, 208
Price, J. L., 63, 209
Price, L. A., 27, 202
Prieto, G., 36, 201
Prinzel, L. J. I., 35, 208

-
- Quaiser-Pohl, C., 49, 50, 62, 208
Quinn, D. M., 15, 127, 145, 199, 201, 210
- Ratcliff, R., 161, 191, 192, 208
Raven, J., 120, 208
Raven, J. C., 120, 208
Redersdorff, S., 129, 206, 208
Regional Educational Laboratory Southeast, 59, 198
Rettinger, D. A., 30, 207
Revelle, W., 150, 204
Régner, I., 56, 58, 61, 63, 80, 113, 118, 203, 204, 208
Ribble, C., 120, 211
Richardson, C., 207
Richardson, J. T. E., 31, 158, 208
Richeson, J. A., 187, 205
Rilea, S. L., 41, 42, 208
Rimrodt, S. L., 32, 200
Ripley, A., 189, 208
Robert, M., 16, 30, 118, 127, 128, 184, 208
Roberts, A. E., 73, 211
Roberts, J. E., 42, 208
Rodgers, M. A., 36, 211
Roger, C., 195, 199, 208
Rosabianca, A., 16, 199
Rosenthal, H. E. S., 11, 203
Rosenthal, R., 54, 208
Roskos-Ewoldsen, B., 41, 208
Routtenberg, A., 148, 208
Rumiati, R. I., 43, 210
Ryan, A. M., 109, 110, 207
Rydell, R. J., 60, 146, 198
- Saito, T., 31, 209
Sanchez, D., 59, 209
Sanchez-Mazas, M., 191, 199
Sanders, A. F., 17, 144, 147–153, 159, 164, 166, 178, 180, 186, 188, 194, 209
Sarrazin, P., 56, 200
Saunders, K. A., 36, 211
Savoie, N., 30, 208
Scali, R. M., 62, 209
Schadow, J., 33, 204
Schadron, G., 53, 205
Schimel, J., 11, 59, 206, 209
Schmader, T., 17, 55, 59–62, 72, 89, 143, 146, 147, 151, 178, 188, 204, 209
Schmitzer-Torbert, N., 33, 209
Schmitz-Williams, I. C., 34, 201
Schon, K., 206
- Schoofs, D., 11, 203
Schultz, W., 12, 200
Sekaquaptewa, D., 59, 63, 140, 205, 209
Senf, W., 73, 202
Shah, P., 28, 30, 207, 209
Shapiro, J., 55, 209
Sharps, M. J., 63–66, 69–72, 74, 78, 80, 209
Shepard, R. N., 27, 65, 68, 153, 154, 156, 178, 209
Shih, M., 56, 59, 70, 80, 198, 209
Shiina, K., 31, 209
Siedentopf, C. M., 211
Siegel-Hinson, R. I., 42, 49, 209
Signorella, M. L., 48, 49, 53, 62, 118, 126, 185, 209
Silverman, I., 29, 38, 202, 210
Sjomeling, M., 56, 210
Slabbekoorn, D., 12, 203
Smeding, A., 208
Smievoll, A. I., 210
Smith, W., 34, 201
Spencer, S. J., 15, 56, 58–60, 75, 77, 79, 81, 82, 115, 127, 145, 181, 183, 184, 199, 201, 210
Springer, S. P., 40, 210
Staffen, W., 44, 211
Stafford, R. E., 120, 210
Steele, C. M., 11, 13–16, 52–60, 63, 66, 75, 78, 82, 117, 144, 145, 180, 181, 198, 199, 210
Steele, J., 59, 198
Stern, C. E., 206
Sternberg, S., 148, 210
Stiner, M. C., 39, 205
Stone, J., 56, 60, 145, 200, 210
Stricker, L. J., 110, 210
Strong, R. A., 33, 201
Strupp, J. P., 210
Stumpf, H., 35, 210
Sullivan, K., 11, 212
Sutherland, R. J., 33, 198
Suzuki, K., 31, 209
Swisher, J. D., 206
- Tagaris, G. A., 43, 210
Takahira, S., 48, 207
Takeuchi, Y., 48, 207
Tan, U., 48, 114, 203
Taylor, D. G., 49, 207
Testart, A., 39, 210

-
- Thinus-Blanc, C., 208
Thompson, M., 59, 140, 209
Thomsen, T., 33, 43, 44, 192, 210
Thomson, W. L., 200
Thurstone, L. L., 27, 210
Thurstone, T. G., 27, 210
Tobler, P. N., 12, 200
Tomasino, B., 43, 210
Tooby, J., 29, 39, 210, 211
Tricarico, M. D., 206
Tucker, D. M., 188, 201

Uban, K. A., 202
Ugurbil, K., 210
Unterrainer, J., 44, 211

Valian, V., 191, 211
Vandenberg, S. G., 26–28, 33, 64, 65, 67, 76,
83, 84, 86, 120, 153, 154, 182, 211
Van Goozen, H. M., 12, 203
Van Zandt, T., 165, 211
Vasta, R., 120, 211
Vecchi, T., 29, 31, 62, 158, 200, 208
Veen, V. van, 188, 195, 211
Vidal, F., 195, 199, 208
Vincent, S. B., 192, 211
Vogel, D. S., 42, 211
Vogel, J. J., 42, 192, 211
Voyer, D., 11, 12, 28, 29, 36, 41, 62, 83, 99,
127, 182, 199, 211

Voyer, S., 11, 211

Walter, K. D., 73, 211
Walton, G. M., 55, 74, 211
Wanke, I., 73, 202
Ward, W. C., 110, 210
Webb, R. M., 20, 211
Weiss, E., 12, 33, 44, 73, 192, 211
Welsh, R. C., 201
Welton, A. L., 63, 209
Wheatley, G. H., 31, 200
Wheeler, S. C., 55, 60, 109, 184, 212
Williams, J. K., 209
Williams, K. J., 55, 203
Wilson, W. L., 202
Witt, S. D., 47, 212
Wolford, J. L., 50, 201
Wraga, M., 11, 12, 14, 45, 68–70, 72–75, 78,
79, 177, 183, 187, 193, 212
Wranek, U., 44, 211
Wright, T. M., 29, 62, 203
Wüstenberg, T., 12, 33, 204

Yagi, A., 31, 156, 204
Yerkes, R. M., 150, 212
Yzerbyt, V., 53, 205

Zaiyouna, R., 207
Zhu, Y.-S., 199

Index des sujets

- Analyse distributionnelle, 153, **164–166**, 178, 191
Arousal, 146, 148, 149, 151
Attentes à l'égard du test, 45, 64, 67, 122, 126
Auto-régulation, voir Contrôle cognitif
Auto-évaluation, 104–105, 112, 123, 125, 137 comparative, 105–106, 128
Capacités attentionnelles, voir Ressources cognitives
Charge cognitive, voir Charge mentale
Charge mentale, 31, 34, 60, 111, **145**, **187**
Contrôle cognitif, **146**, 150, 164, 187, 195
Cortex cingulaire antérieur (CCA), 45, 73, 195
Difficulté
 du test, 27–28, 58, 62, 65, 96, 144
 perçue, 103–104, 112, 138
Différences de sexe
 en mathématiques, **19–21**, 49, 56, 61, 190
 verbales, 21–22
 visuo-spatiales, 23–26
Distribution de performance, **81**, 90, 91, 109, 132, 182
Dynamique de la performance, voir Décours de la performance
Décours de la performance, 93–94, 96–97, 135
Échange vitesse / précision, voir Stratégies de sélection de la réponse
Effort, 146, 148, **149–150**, 151, 164, 187, 194
Électroencéphalographie (EEG), 42, 194
Enjeu du résultat du test, **58**, 141, 144, 187
Évolutionniste (hypothèse), 38–40, 46
Ex-gaussienne (distribution), 165–166
Expérience spatiale, 46, 49–50
Falsification (induction), 56, 75, 184
Hormones sexuelles, 39
Identification au domaine de compétence, **57**, 62
Identité sociale, 57, 70
Idéomotrice (hypothèse), 184
Imagerie par résonance magnétique fonctionnelle (IRMf), 43–44, 72–74, 187, 194
Induction, 58, 59, 75
Interférences, 60, 146
Internalisation du stéréotype, **48**, 52
Intervention, voir Induction
Latéralisation fonctionnelle, **40–43**, 50, 193
Lobe
 frontal, 43–45
 pariétal, 43, 44
Modèles bio-psycho-sociaux, **37**, 189
Mémoire de travail, **28–31**, 31, 45, 62, **146**, 151, 187
Mémoire spatiale, 29, 38
Navigation, voir Orientation
Niveau de performance, **58**, 62, **82**, 89, **91**
Niveau de veille, voir Arousal
Orientation, 25, 32, 33, 118
Pensées négatives, 60, **61**, 89, 111, 146
Perception spatiale, 25
Ratio
 de sexe, voir Sexe-ratio
 essais tentés / réussis, 35, 36

Ressources cognitives, 29, 32, 60, 62, 146,
148, 178, 187

Rotation mentale, **26–27**
test, **26**, 27, 28, **83–84**, 123, 127
informatisé, 156

Ruminations, voir Pensées négatives

Section du test, 84, 89, 96, 111, 135

Sexe-ratio, 88, **90–93**

Signification du test, voir Attentes à l'égard
du test

Socialisation, 38, **46–47**, 50, 115, 189

Stratégies
de sélection de la réponse, **34–37**, 175
de traitement de l'information, **31–34**,
43, 192
analogique, **26**, 32
analytique, **32**, 34, 38

Stéréotype, **53–54**
adhésion au, 58, 85, 101–102, 112, 136
connaissance du, 57, 84, 100–101, 112
de sexe, 47, 48, 53, 61
interférence (menace) du, **54**
questionnaire, 84–85, 108, 112–114,
120–122
stereotype lift, 74, 88–89, 94, 110
visuo-spatial, 62, 118, 126–128

Taille d'effet, 22, 23, 94–95, 99

Traitement de l'information
actif, 29–31
séquentiel, 148
vitesse, 27

Visualisation spatiale, 24

Annexes

Annexe A

Étude 1

MENTAL ROTATION TEST (MRT) : INSTRUCTIONS

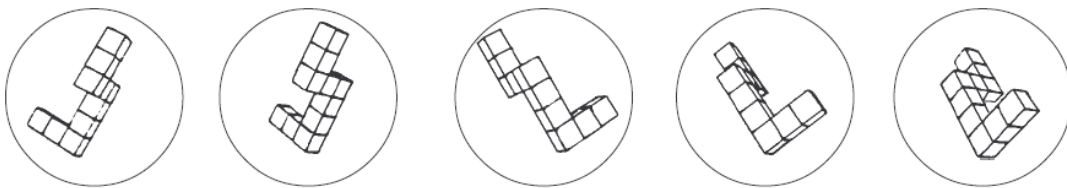
Date de naissance : _____

SEXE : _____

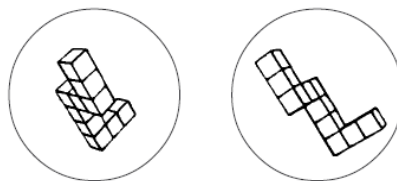
L'exercice qui t'est proposé est un test.

Ce test permet de savoir si tu es capable de reconnaître un objet alors qu'il est mélangé avec d'autres objets qui lui ressemblent.

Par exemple, voici ci-dessous **cinq dessins du même objet, mais vu sous des angles différents**. Observe attentivement chaque dessin pour te rendre compte qu'il s'agit bien à chaque fois du même objet :

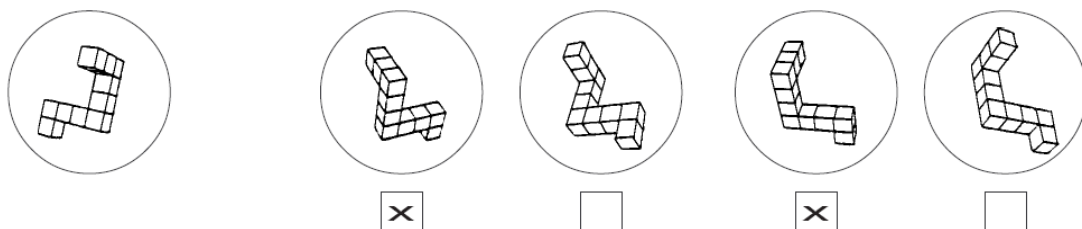


Par contre ci-dessous, voici deux dessins d'un **objet différent** de celui que tu viens de voir. Observe et compare ces dessins ci-dessous pour bien te rendre compte qu'il est différent du précédent :



Pour chaque problème, on te montre d'abord à gauche l'objet que tu dois retrouver. Tu dois ensuite indiquer parmi les quatre objets de droite ceux qui sont identiques à l'objet que tu dois retrouver.

A chaque fois il y a deux réponses justes. Mets des croix (X) dans les cases sous les dessins corrects. Voici un exemple de réponse :

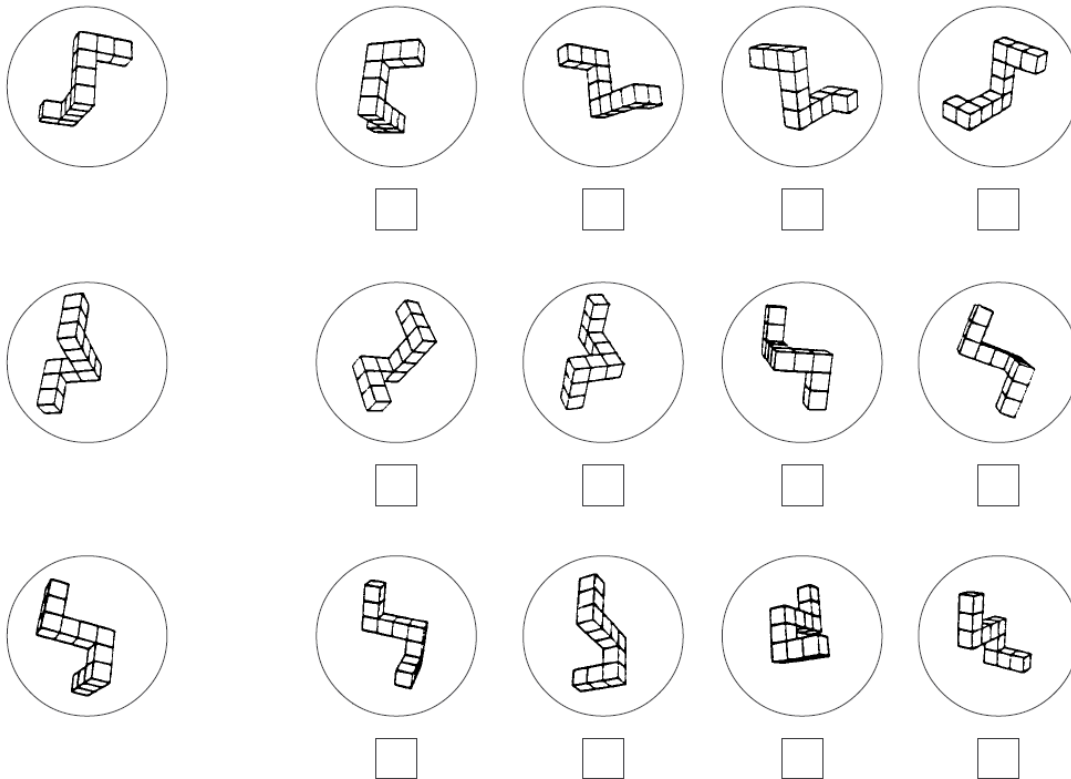


Tournez la page

MENTAL ROTATION TEST (MRT) : INSTRUCTIONS

Maintenant voici trois problèmes pour t'entraîner.

N'oublie pas qu'à chaque fois il y a **deux dessins corrects** parmi les quatre proposés.



Réponses :

1 – premier et second dessins corrects

2 – premier et troisième dessins corrects

3 – deuxième et troisième dessins corrects

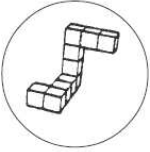
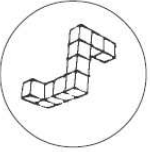
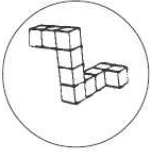
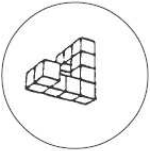
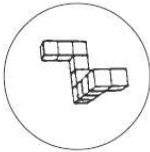
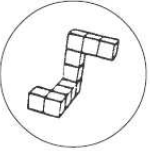
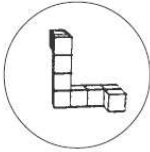
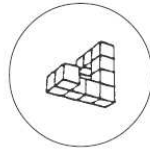
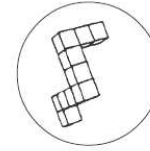
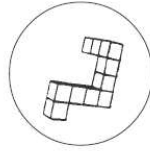
Une dernière chose : ce test est en deux parties. Tu as 3 minutes pour faire chaque partie. Il faut absolument respecter ce délai. Chaque partie vous sera distribuée face contre table. **Tu ne retourneras les feuilles pour commencer à répondre au test seulement lorsque l'on te le demandera.**

Réponds aussi vite que tu peux, mais en essayant de faire juste à chaque fois. Ne réponds surtout pas au hasard.

Dans la condition d'activation la phrase : « Pour information, en général, sur ce test, les garçons et les filles n'obtiennent pas les mêmes résultats ; ils n'ont pas la même performance. » était ajoutée.

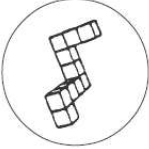
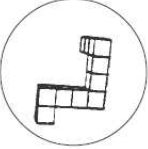
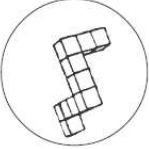

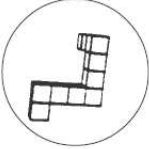
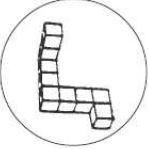
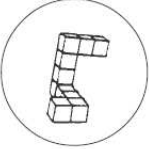
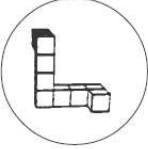
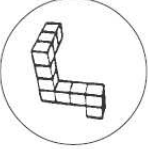
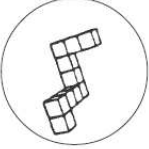
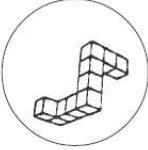
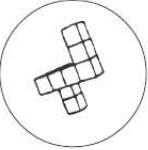
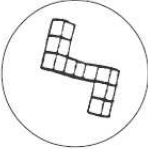
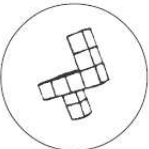
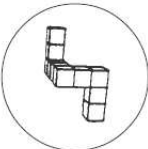
Dans la condition d'activation la phrase : « Pour information, en général, sur ce test, les garçons et les filles obtiennent les mêmes résultats ; ils ont la même performance. » était ajoutée.

MENTAL ROTATION TEST (MRT) : SECTION 1

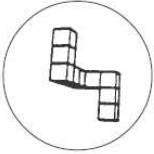
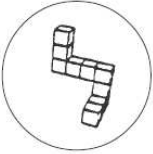
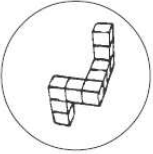
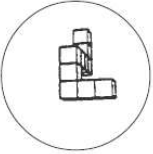
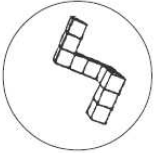
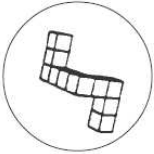
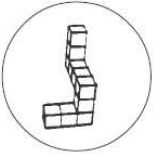
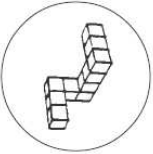
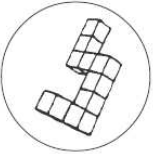
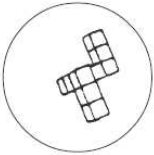
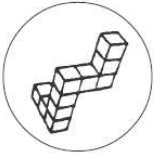
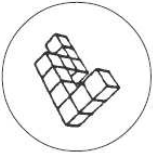
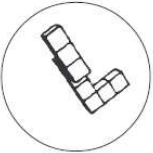
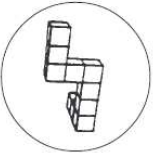
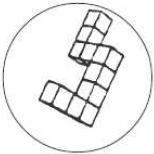
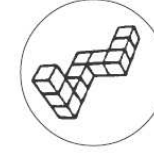
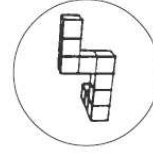
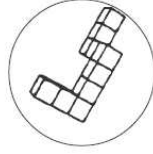
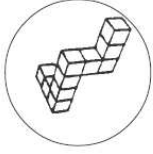
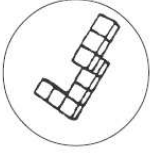
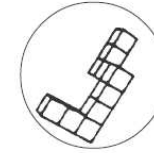
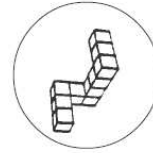
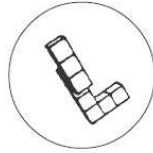
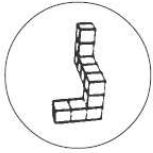
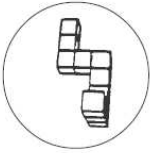
1					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

TOURNER LA PAGE

MENTAL ROTATION TEST (MRT) : SECTION 1

6		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
7		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
8		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
9		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
10		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>

MENTAL ROTATION TEST (MRT) : SECTION 2

11		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
12		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
13		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
14		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
15		 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>

TOURNER LA PAGE

MENTAL ROTATION TEST (MRT) : SECTION 2

16					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
17					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
18					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
19					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
20					
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

QUESTIONNAIRE : INSTRUCTIONS

Date de naissance : _____

SEXE : _____

Dans ce questionnaire, tu devras indiquer pour chaque question si tu es d'accord ou pas avec la phrase proposée. Mais attention ce questionnaire est en deux parties :

Dans la première partie tu dois répondre *en fonction de ce que les gens pensent*, et non de ce que tu penses toi.

Voici un exemple de question :

Y a-t-il des gens qui pensent que les cours d'anglais sont plus importants que les cours de musique ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

Tu vois que pour répondre, il faut cocher le chiffre qui correspond le mieux à ta réponse. Il n'y a pas de bonnes ou de mauvaises réponses. Seul ton avis est important.

Dans la seconde partie par contre tu devras répondre aux mêmes questions, mais cette fois *en répondant en fonction de ce que tu penses toi-même, de ta propre opinion*.

Personnellement, je pense que les cours d'anglais sont plus importants que les cours de musique.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

Les questions auxquelles tu devras répondre dans ce questionnaire sont liées aux capacités des femmes et des hommes, dans différentes situations.

QUESTIONNAIRE : CONNAISSANCE DU STÉRÉOTYPE

G1

Y a-t-il des gens qui pensent que les femmes ont plus de mal que les hommes pour se servir d'une carte routière ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

G2

Y a-t-il des gens qui pensent que, en randonnée, les femmes ont plus de mal que les hommes pour retrouver leur chemin ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

G3

Y a-t-il des gens qui pensent que les femmes ont plus de mal que les hommes pour recopier parfaitement un dessin compliqué ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

G4

Y a-t-il des gens qui pensent que lorsqu'elles jouent au basket, les femmes se placent moins bien que les hommes sur le terrain pour recevoir et envoyer le ballon ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

QUESTIONNAIRE : CONNAISSANCE DU STÉRÉOTYPE

G5

Y a-t-il des gens qui pensent que les femmes ont plus de mal que les hommes pour imaginer un objet en trois dimensions (boîte, meuble, maison...) à partir d'un dessin en deux dimensions ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

G6

Y a-t-il des gens qui pensent que les femmes ont plus de difficultés que les hommes pour reconnaître un endroit où elles sont déjà passées.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Personne	Quelques uns	Une personne sur deux	Beaucoup	Tout le monde

QUESTIONNAIRE : ADHÉSION AU STÉRÉOTYPE

P1

Personnellement, je pense que les femmes ont plus de mal que les hommes pour se servir d'une carte routière.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

P2

Personnellement, je pense que, en randonnée, les femmes ont plus de mal que les hommes pour retrouver leur chemin.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

P3

Personnellement, je pense que les femmes ont plus de mal que les hommes pour recopier parfaitement un dessin compliqué.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

P4

Personnellement, je pense que lorsqu'elles jouent au basket, les femmes se placent moins bien que les hommes sur le terrain pour recevoir et envoyer le ballon.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

QUESTIONNAIRE : ADHÉSION AU STÉRÉOTYPE

P5

Personnellement, je pense que les femmes ont plus de mal que les hommes pour imaginer un objet en trois dimensions (boîte, meuble, maison...), à partir d'un dessin en deux dimensions.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

P9

Personnellement, je pense que les femmes ont plus de difficultés que les hommes pour reconnaître un endroit où elles sont déjà passées.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Pas du tout d'accord	Plutôt pas d'accord	Ni d'accord, ni pas d'accord	Plutôt d'accord	Tout à fait d'accord

QUESTIONNAIRE : PERCEPTION DU MRT

T1

Pour toi, ce test était :

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très difficile	Difficile	Ni facile, ni difficile	Facile	Très facile

T2

Comment te juges-tu sur ce test :

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très mauvais(e)	Mauvais(e)	Ni bon(ne), ni mauvais(e)	Bon(ne)	Très bon(ne)

T3

Sur ce test, comment te juges-tu par rapport aux autres élèves de la classe :

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Beaucoup moins bon(ne)	Moins bon(ne)	Comme eux	Meilleur(e)	Bien meilleur(e)

T4

Selon toi, comparativement aux garçons, les filles réussissent ce test :

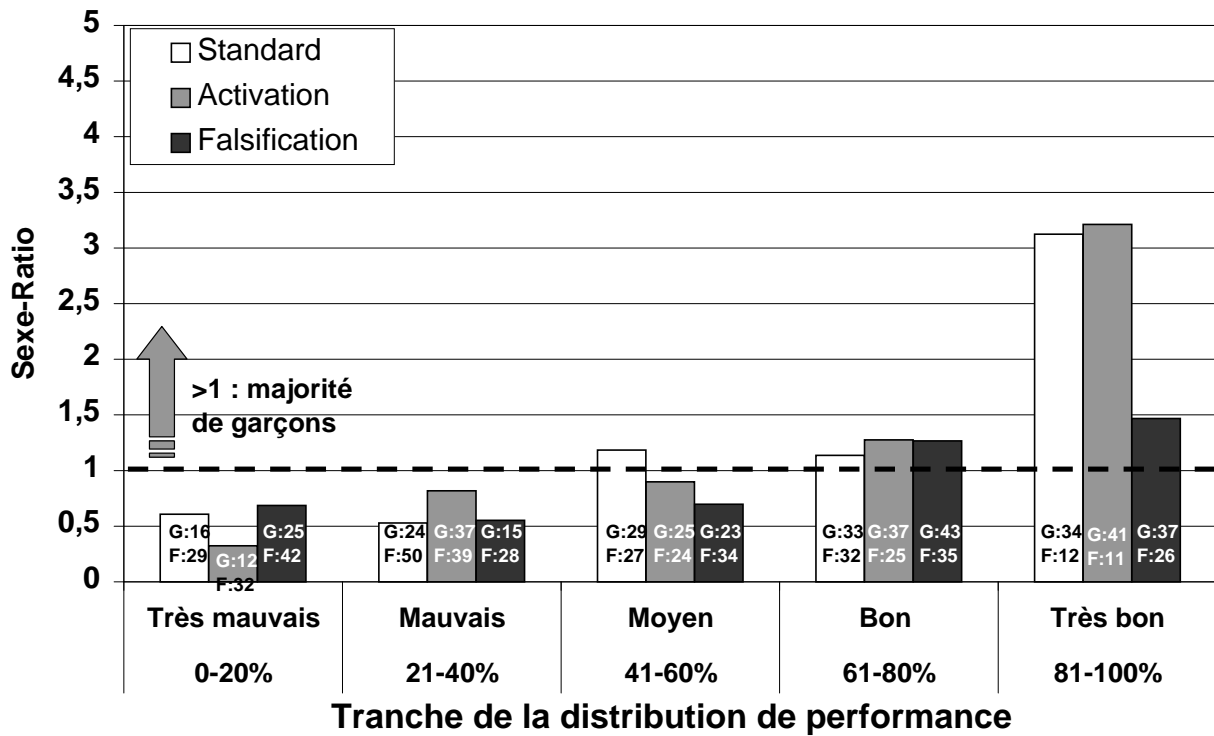
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Beaucoup moins bien	Moins bien	Comme eux	Mieux	Beaucoup mieux

T5

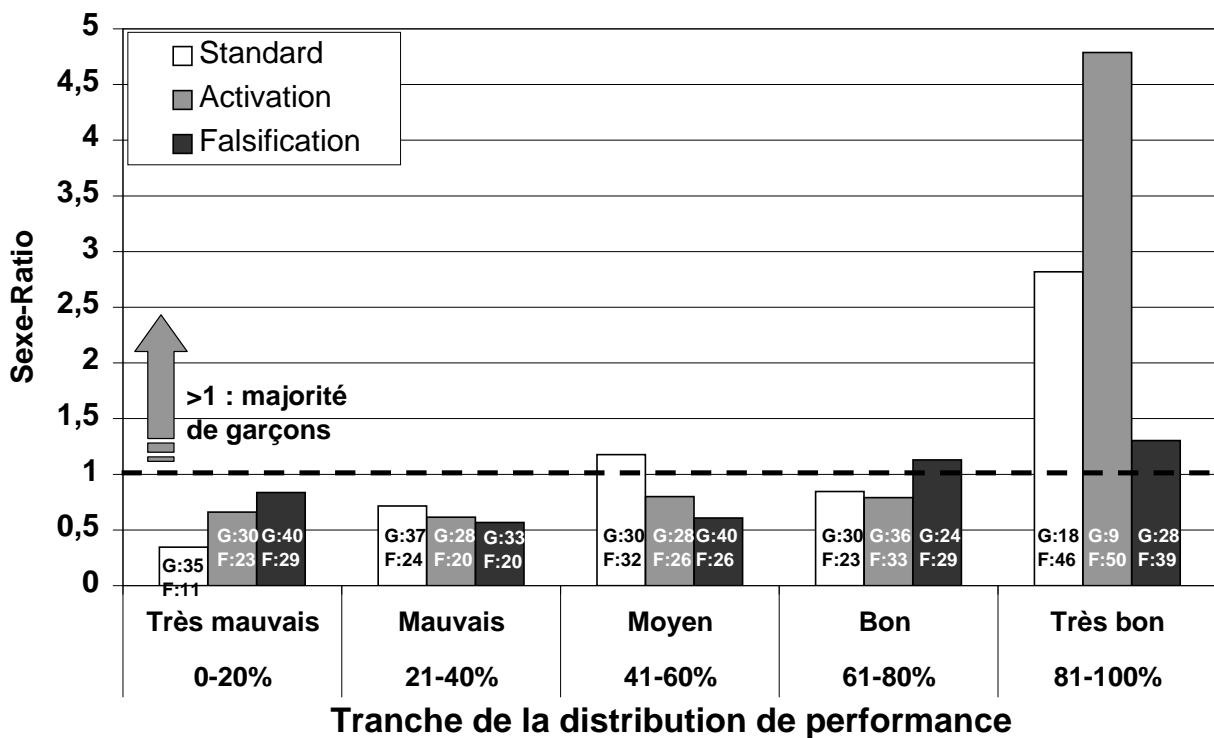
Selon toi, comparativement aux personnes âgées, les jeunes réussissent ce test :

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Beaucoup moins bien	Moins bien	Pareil	Mieux	Beaucoup mieux

SCORES ÉCHANTILLON TOTAL



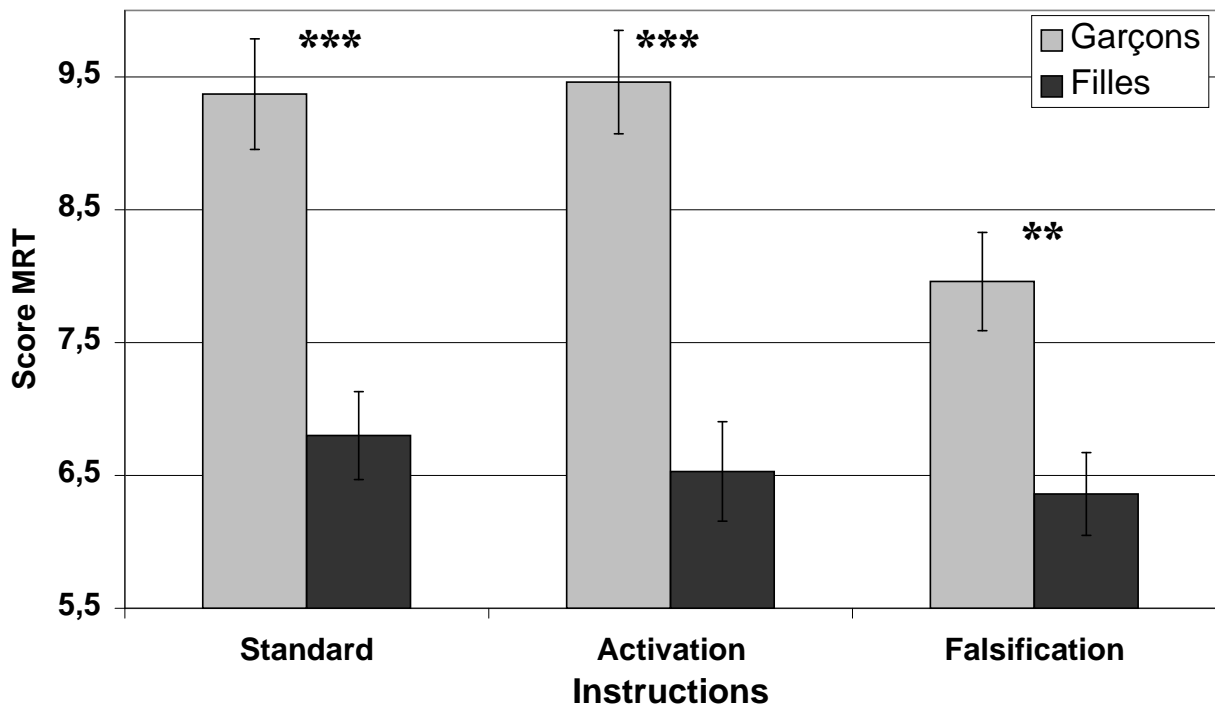
(a) Première partie du MRT



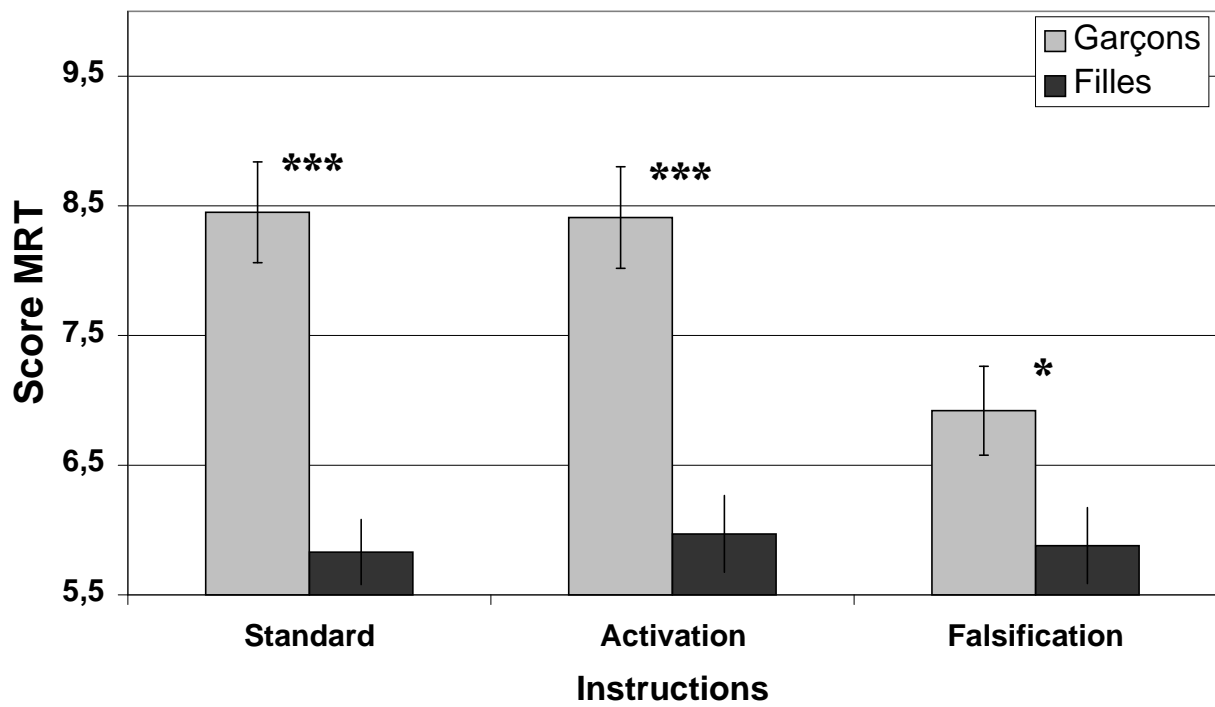
(b) Seconde partie du MRT

FIGURE A.1 — Sexe-ratios pour différentes tranches de performance. Chaque tranche représente 20% de l'effectif total..

SCORES ÉCHANTILLON TOTAL



(a) Première partie du MRT



(b) Seconde partie du MRT

*** : $p < .001$; ** : $p < .01$; * : $p < .05$

FIGURE A.2 – Interaction *Sexe* \times *Instructions* \times *Section* (échantillon total).

Annexe B

Étude 2

Date de naissance : _____

Sexe : _____

Dans ce questionnaire, nous vous présentons des exemples de tests. Vous devez indiquer pour chacun de ces tests quels résultats, selon vous, est obtenu par certaines catégories de personnes.

Voici un exemple de question :

Dans ce test, il faut arriver à souligner le plus vite possible sans en oublier une seule, chaque lettre « A » contenue dans un texte.

Exemple :

Ce dernier souligne que l'on ne peut jamais décider de la valeur de vérité d'une règle induite tant qu'elle est seulement confirmée par des exemples, alors qu'un seul contre-exemple falsifiant permet de décider de sa fausseté de manière valide. Par exemple après avoir observé plusieurs cygnes blancs on peut émettre l'hypothèse : Tous les cygnes sont blancs. Peut-on affirmer que cette règle est vraie ? Que cette règle est fausse ?

Réponse :

Ce dernier souligne que l'on ne peut jamais décider de la valeur de vérité d'une règle induite tant qu'elle est seulement confirmée par des exemples, alors qu'un seul contre-exemple falsifiant permet de décider de sa fausseté de manière valide. Par exemple après avoir observé plusieurs cygnes blancs on peut émettre l'hypothèse : Tous les cygnes sont blancs. Peut-on affirmer que cette règle est vraie ? Que cette règle est fausse ?

Selon vous, sur ce test, le résultat des personnes âgées est en général :

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

N'oubliez pas que dans ce questionnaire, ce qui nous intéresse, c'est la première réponse qui vous vient à l'esprit. Il n'y a pas de bonne ou de mauvaise réponse. Répondez le plus spontanément possible, sans perdre de temps à réfléchir.

Dans ce test, il faut trouver le plus vite possible le plus de mots qui commencent par une lettre donnée.

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :

K_____

Réponses :

Koala ; Kilogramme ; Kaki ; Képi ; Kiwi ; Krach ; Klaxon ; Kidnapper ; ...

Selon vous, sur ce test, le résultat des personnes jeunes est en général :

FV1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

FV2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

FV3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

FV4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général :

FV5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général : FV6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général : FV7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général : FV8

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

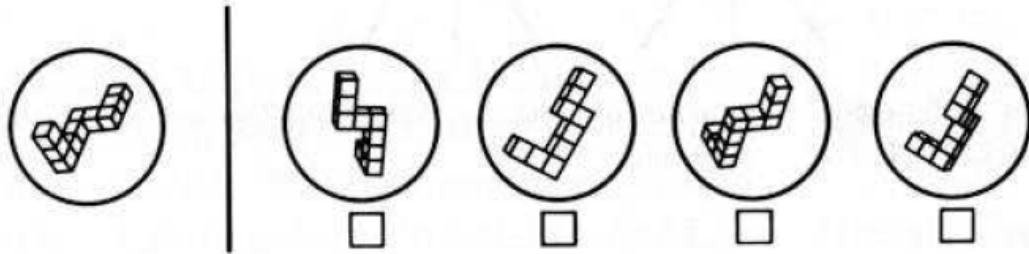
En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat : FV9

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

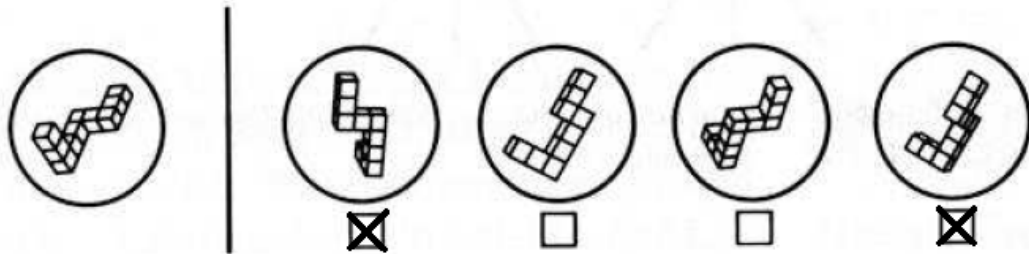
QUESTIONNAIRE - ATTENTES ENVERS DIFFÉRENTS TESTS

Dans ce test, on vous demande de trouver parmi les 4 figures dessinées à droite les 2 qui sont des représentations correctes de l'objet présenté à gauche, mais vu sous des angles différents.

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :



Réponse :



Selon vous, sur ce test, le résultat des personnes jeunes est en général :

RM1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

RM2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

RM3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

RM4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général :

RM5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général :

RM6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général :

RM7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général :

RM8

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat :

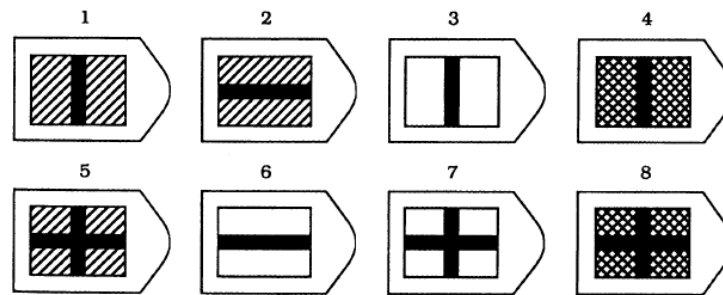
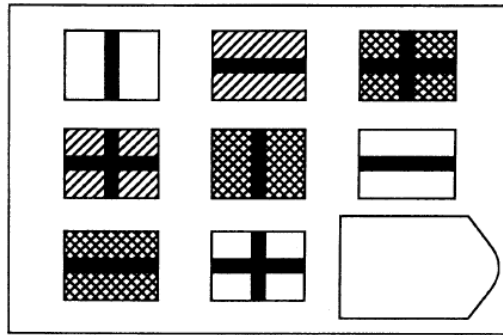
FV9

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

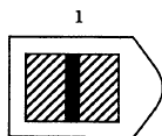
QUESTIONNNAIRE - ATTENTES ENVERS DIFFÉRENTS TESTS

Dans ce test, il faut trouver quelle est la pièce manquante du puzzle. Cette pièce n'est pas choisie au hasard. Pour pouvoir la trouver, il faut savoir que le puzzle est construit en suivant une règle. Si vous trouvez la règle qui s'applique dans les deux premières lignes et les deux premières colonnes, vous saurez quelle est la pièce manquante.

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :



Réponse :



Selon vous, sur ce test, le résultat des personnes jeunes est en général :

MP1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

MP2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

MP3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

MP4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général :

MP5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général :

MP6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général :

MP7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général :

MP8

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat :

FV9

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Dans ce test, on doit écouter une liste de mot qu'il faut retenir pour les rappeler 20mn plus tard :

Exemple de mots à apprendre :

Livre ; Symbole ; Inquiet ; Sommeil ; Justice ; Autoroute ;
Argent ; Diamant ; Enfant ; Hôpital ; Film ; Lampe

Selon vous, sur ce test, le résultat des personnes jeunes est en général :

MV1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

MV2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

MV3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

MV4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général :

MV5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général : MV6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général : MV7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général : MV8

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

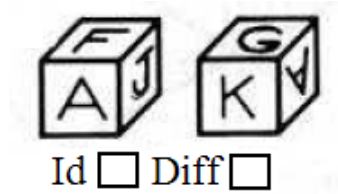
En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat : FV9

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

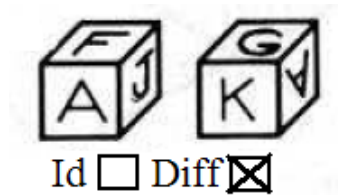
QUESTIONNAIRE - ATTENTES ENVERS DIFFÉRENTS TESTS

Dans ce test, il faut indiquer si deux cubes posés côte à côte peuvent être identiques (dans ce cas il faut cocher la case « Id ») ou différents (dans ce cas il faut cocher la case « Diff »).

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :



Réponse :



Selon vous, sur ce test, le résultat des personnes jeunes est en général :

IB1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

IB2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

IB3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

IB4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général : IB5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général : IB6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général : IB7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général : IB8

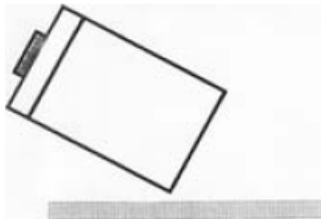
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat : FV9

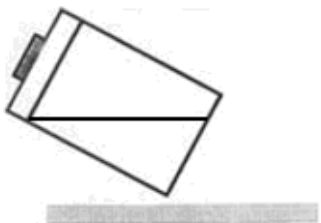
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Dans ce test, il faut imaginer que le flacon dessiné ci-dessous est à moitié rempli d'eau. Pour répondre il faut tracer un trait au crayon qui indique le niveau de l'eau dans le flacon.

Exemple (en fait dans le test il y a plusieurs problèmes comme celui-ci) :



Réponse :



Selon vous, sur ce test, le résultat des personnes jeunes est en général :

WL1

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des personnes âgées est en général :

WL2

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des femmes sur ce test est en général :

WL3

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, le résultat des hommes sur ce test est en général :

WL4

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en géométrie est en général :

WL5

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en géométrie est en général :

WL6

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves forts en sciences est en général :

WL7

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Selon vous, sur ce test, le résultat des élèves faibles en sciences est en général :

WL8

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

En ce qui vous concerne, vous pensez que sur ce test vous obtiendriez un résultat :

FV9

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7
Très mauvais	Mauvais	Plutôt mauvais	Ni bon ni mauvais	Plutôt bon	Bon	Très bon

Annexe C

Étude 3

QUESTIONNAIRE - POST-TEST MRT

SE1

En général, les hommes sont supérieurs aux femmes concernant les 'compétences spatiales' (capacité à s'orienter dans l'espace, à lire des cartes routières, à manipuler mentalement des objets en 3 dimensions).

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
Absolument pas d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Absolument d'accord

SE2

Il n'y a pas de différences réelles entre les hommes et les femmes concernant les 'compétences spatiales'

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
Absolument pas d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Absolument d'accord

SE3

Il est possible qu'en matière de compétences spatiales, les femmes aient d'une façon générale moins de facilités que les hommes.

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
Absolument pas d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Absolument d'accord

SE4

Les femmes paraissent plus douées que les hommes dans la plupart des domaines qui réclament des compétences spatiales

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
Absolument pas d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Absolument d'accord

SE

Comment jugez-vous votre performance sur ce test (avec les figures) ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Médiocre	Plutôt mauvaise	Acceptable	Plutôt bonne	Excellente

QUESTIONNAIRE - POST-TEST MRT

CSE1

Par rapport à la performance des personnes de même sexe, de même filière et de même niveau d'étude que vous, comment situeriez-vous votre propre performance sur ce test ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très inférieure	Légèrement inférieure	Similaire	Légèrement supérieure	Très supérieure

CSE2

Par rapport à la performance des personnes *de sexe différent*, de même filière et de même niveau d'étude que vous, comment situeriez-vous votre propre performance sur la première série ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très inférieure	Légèrement inférieure	Similaire	Légèrement supérieure	Très supérieure

CSE3

Par rapport à la performance des personnes *âgées (>70 ans)*, comment situeriez-vous votre propre performance sur ce test ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très inférieure	Légèrement inférieure	Similaire	Légèrement supérieure	Très supérieure

MC1

Sur ce test, comment situeriez-vous la performance des participants de même sexe que vous par rapport à celle des participants de sexe opposé ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très inférieure	Légèrement inférieure	Similaire	Légèrement supérieure	Très supérieure

MC2

Sur ce test, comment situeriez-vous la performance des personnes de votre âge par rapport à celle des personnes âgées ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très inférieure	Légèrement inférieure	Similaire	Légèrement supérieure	Très supérieure

TPD

Que diriez-vous du niveau de difficulté de la première série d'exercices ?

<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Très facile	Plutôt facile	Entre les deux	Plutôt difficile	Très difficile

RÉSUMÉ

Les différences de sexe observées dans les tests de rotation mentale 3D sont connues pour être typiquement fortes et robustes, ancrées dans des mécanismes biologiques, et amplifiées au cours de la socialisation via l'intégration des rôles sexués. Ce consensus, cependant, ne prend pas en compte différents mécanismes liés à l'intervention de stéréotypes négatifs concernant les femmes dans le domaine visuo-spatial et susceptibles d'altérer leurs performances au cours même du test. Nous soutenons que ces mécanismes sont importants, sinon cruciaux, pour mieux comprendre l'origine des différences en question. Adossés notamment à la littérature sur la « menace du stéréotype », nos résultats expérimentaux (4 études, 1500 participants) nous conduisent à considérer la malléabilité de ces différences et leur ancrage dans des constructions culturelles (stéréotypes sociaux). Notre recherche contribue à ce titre à l'étude de la régulation sociale des processus cognitifs de base.

MOTS CLEFS

Activités visuo-spatiales, Différence de sexe, Régulation sociale, Rotation mentale, Stéréotypes sociaux

ABSTRACT

Social stereotypes and regulation of sex difference in 3D mental rotations : An experimental perspective.

The sex differences associated with 3D mental rotation tests are typically thought to be especially strong, robust, rooted in biological mechanisms, and amplified by gender role socialization. This consensus, however, does not take into account several mechanisms related to the intervention of negative stereotypes about women in the visuo-spatial domain, which may interfere with task performance. Here, we claim that these mechanisms are important otherwise crucial to understand better the origin of the differences in question. Based on the stereotype threat literature, our experimental findings (4 studies, 1500 participants) lead to consider the malleability of these differences and their anchoring in cultural knowledge (social stereotypes). As such, our research contributes to the study of the social regulation of basic operations of cognition.

KEYWORDS

Mental rotation, Sex difference, Social regulation, Social stereotype, Visuo-spatial activities

FORMATION DOCTORALE

Psychologie

Laboratoire de psychologie cognitive

CNRS UMR 6146

Université de Provence, Centre St-Charles

Pôle 3C — Comportement, Cerveau, Cognition — Bâtiment 9, Case D

3, Place Victor Hugo

13331 Marseille Cedex 1 France