



**HAL**  
open science

# Rôle de la mémoire dans la métacognition : Utilisation des informations épisodiques et sémantiques

Méline Devaluez

► **To cite this version:**

Méline Devaluez. Rôle de la mémoire dans la métacognition : Utilisation des informations épisodiques et sémantiques. Psychologie. Université Grenoble Alpes [2020-..], 2023. Français. NNT : 2023GRALS010 . tel-04433511

**HAL Id: tel-04433511**

**<https://theses.hal.science/tel-04433511v1>**

Submitted on 2 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : ISCE - Ingénierie pour la Santé la Cognition et l'Environnement

Spécialité : PCN - Sciences cognitives, psychologie et neurocognition

Unité de recherche : Laboratoire de Psychologie et NeuroCognition

## Rôle de la mémoire dans la métacognition : Utilisation des informations épisodiques et sémantiques

### The role of memory in metacognition: Use of episodic and semantic information

Présentée par :

**Méline DEVALUEZ**

#### Direction de thèse :

**Céline SOUCHAY**

DIRECTEUR DE RECHERCHE, CNRS, Université Grenoble Alpes

Directrice de thèse

**Christopher MOULIN**

PROFESSEUR DES UNIVERSITÉS, Université Grenoble Alpes

Co-Directeur de thèse

#### Rapporteurs :

**Marie IZAUTE**

PROFESSEUR DES UNIVERSITÉS, Université Clermont-Auvergne

**Louis RENOULT**

PROFESSEUR ASSOCIÉ, University of East Anglia, Royaume-Uni

#### Thèse soutenue publiquement le **20 mars 2023**, devant le jury composé de :

**Céline SOUCHAY**

DIRECTEUR DE RECHERCHE, CNRS, Université Grenoble Alpes

Directrice de thèse

**Marie IZAUTE**

PROFESSEUR DES UNIVERSITÉS, Université Clermont-Auvergne

Rapporteuse,  
Présidente du jury

**Louis RENOULT**

PROFESSEUR ASSOCIÉ, University of East Anglia, Royaume-Uni

Rapporteur

**Christine BASTIN**

CHARGÉE DE RECHERCHE, Liège Université, Belgique

Examinatrice

**Matthias KLIEGEL**

PROFESSEUR, Université de Genève, Suisse

Examineur

**Marcela PERRONE-BERTOLOTTI**

MAÎTRE DE CONFÉRENCES, Université Grenoble Alpes

Examinatrice

#### Invités :

**Christopher MOULIN**

PROFESSEUR DES UNIVERSITÉS, Université Grenoble Alpes

Co-Directeur de thèse





# Résumé

---

Une partie du champ de la métacognition de la mémoire a été façonnée par la distinction épisodique/sémantique proposée par Tulving (1972). Par exemple, des preuves expérimentales et neuropsychologiques suggèrent que des processus distincts sont impliqués dans certains paradigmes de métacognition portant sur la mémoire sémantique et ceux portant sur la mémoire épisodique. De plus, la relation entre performance à la tâche de mémoire et jugements métacognitifs est un débat central dans la littérature. Néanmoins, en parallèle de cette distinction épisodique/sémantique, de nombreux travaux suggèrent que mémoires sémantique et épisodique ne fonctionnent pas de manière totalement indépendante, notamment chez les sujets sains. Plus particulièrement, plusieurs études montrent l'importance et l'influence des traitements sémantiques dans l'encodage et la récupération en mémoire épisodique mais aussi dans la capacité des personnes à auto-évaluer leurs performances sur ces tâches de mémoire épisodique. Ainsi, le but de cette thèse était d'approfondir les connaissances sur la relation entre processus mnésiques et jugements métacognitifs. Deux points de cette relation ont été abordés. Tout d'abord, nous nous sommes intéressés à la relation spécifique entre type de mémoire mis en jeu dans la tâche de premier ordre (i.e., épisodique ou sémantique) et performances métacognitives (Chapitres 4 et 5). Un second objectif était d'approfondir la compréhension de l'influence des traitements sémantiques dans les jugements métacognitifs (Chapitres 5, 6, et 7).

Dans un premier temps, nous avons réalisé deux méta-analyses portant sur la distinction entre *feeling-of-knowing* épisodique (eFOK) et sémantique (sFOK) dans le vieillissement sain. Ces méta-analyses ont apporté des preuves supplémentaires à la contribution du déclin en mémoire épisodique dans la diminution de la sensibilité métacognitive du eFOK chez les personnes âgées. Deuxièmement, nous avons évalué les capacités métacognitives (épisodiques et sémantiques) de personnes présentant une pathologie de la mémoire sémantique : la démence sémantique. Les performances de ces personnes ont été comparées à celles de personnes âgées saines et à des personnes atteintes de maladie d'Alzheimer. Cette étude n'a pas permis de mettre en avant de déficit métacognitif sémantique dans cette population, ni de double dissociation entre type de déficit de premier ordre et performances métacognitives.

Ensuite, 4 études ont montré l'importance des traitements sémantiques pour prédire de façon précise les performances sur une tâche de reconnaissance en mémoire épisodique. Enfin, 2 autres études portant sur la fluence conceptuelle ont montré l'influence des traitements sémantiques d'images sur les jugements métacognitifs rétrospectifs et prédictifs en mémoire épisodique. L'ensemble des résultats de cette thèse contribuent à montrer que les processus mnésiques interviennent dans les jugements métacognitifs. Notamment, nous mettons en avant l'importance des capacités de récupération en mémoire épisodique, la familiarité, et l'influence des connaissances sémantiques. De plus, ces processus ne sont pas nécessairement spécifiques au type de mémoire testé sur la tâche de premier ordre puisque la mémoire sémantique intervient également dans des tâches de métacognition épisodique.

**Mots-clés :** Métacognition, mémoire épisodique, mémoire sémantique, démence sémantique

# Abstract

---

Part of the field of memory metacognition has been shaped by the episodic/semantic distinction proposed by Tulving (1972). For example, experimental and neuropsychological evidence suggests that distinct processes are involved in metacognition tasks involving semantic memory and episodic memory. Moreover, the relationship between memory performance and metacognitive judgments is a central debate in the literature. Nevertheless, in parallel to this episodic/semantic distinction, many studies suggest that semantic memory and episodic memory do not function completely independently, especially in healthy subjects. In particular, several studies show the importance and influence of semantic processing in the encoding and retrieval of episodic memories, but also in one's ability to self-assess their performance on episodic memory tasks. Thus, the aim of this thesis was to investigate further the relationship between memory processes and metacognitive judgments. Two aspects of this relationship were addressed. First, we were interested in the specific relationship between the type of memory involved in the first-order task (i.e., episodic or semantic) and metacognitive performance (Chapters 4 and 5). A second objective was to better understand the influence of semantic processing on metacognitive judgments (Chapters 5, 6, and 7).

First, we conducted two meta-analyses on the distinction between episodic and semantic feeling-of-knowing (eFOK and sFOK) in healthy aging. These meta-analyses provided additional evidence for the contribution of episodic memory decline to the decrease in metacognitive sensitivity of eFOK in this population. Secondly, we assessed episodic and semantic metacognitive abilities of individuals with semantic memory impairment: semantic dementia patients. Their performance was compared to that of healthy older adults and to Alzheimer's disease patients. This study did not reveal any particular metacognitive deficit in this population, nor double dissociation between type of first-order deficit and metacognitive performance.

Then, 4 studies showed the importance of semantic processing to accurately predict performance on an episodic memory recognition task. Finally, 2 other studies conducted on the manipulation of conceptual fluency showed the influence of semantic processing of images on retrospective and predictive metacognitive judgments of episodic memory. The results of this thesis show that memory processes are involved in metacognitive judgments. In particular, we highlight the importance of episodic memory retrieval, familiarity, and the influence of conceptual knowledge. Moreover, these types of information are not always specific to the type of memory tested on the first-order task since semantic memory is also involved in episodic metacognition tasks.

**Key words:** Metacognition, episodic memory, semantic memory, semantic dementia

# Remerciements

---

Mes premiers remerciements vont à Christine Bastin, Marie Izaute, Matthias Kliegel, Marcela Perrone-Bertolotti, et Louis Renoult, pour avoir accepté de relire et d'évaluer mon travail.

Je remercie également les membres de mon comité de suivi individuel de thèse Christine Bastin, Nathan Faivre et Hélène Løevenbruck pour leur bienveillance et leurs précieux conseils.

Un très grand merci à mes encadrants. Céline, merci de m'avoir fait confiance pendant 4 ans et demi, et de m'avoir permis de réaliser cette thèse. Merci de m'avoir soutenue, pour ta disponibilité même pendant une pandémie mondiale, et pour tes conseils. Merci aussi pour ton encadrement au-delà du scientifique, ta transmission des connaissances du monde de la recherche. J'aurais beaucoup appris grâce à toi. Chris, merci pour ton aide avec les analyses, de m'avoir permis de donner des cours aussi intéressants, pour ta motivation à aller boire des bières et pour ta bienveillance. Désolée, j'aurais encore besoin de temps pour apprendre à ne pas dire « il y a un problème » ou « mais ». En bref, merci à tous les deux pour votre humanité, qui ont rendu la conduite de cette thèse beaucoup plus facile.

Je remercie également les professionnels ayant contribué à la mise en place du projet META-DEM, sans qui il n'aurait pas vu le jour : Olivier Moreaud, Mathilde Sauvée, Annik Charnallet, Stéphanie Maurice, Serge Belliard, Catherine Merck, et Frédérique Etcharry-Bouyx.

Merci au personnel de la MSH-Alpes et de la plateforme SCREEN, et à Lucile Meunier-Duperray pour son immense travail pour développer « Seniors pour la Science ».

Un très grand merci à tous les participants de mes études pour leur temps, et en particuliers les participants du projet META-DEM pour leur gentillesse.

Mes prochains remerciements vont à Audrey Mazancieux, qui m'a beaucoup apporté lorsque j'étais encore étudiante en Master, mais également pendant ce doctorat. Toujours pleine d'idées ! C'était un plaisir de collaborer avec toi sur la méta-analyse.

Merci également aux étudiantes de Master que j'ai eu le plaisir d'encadrer : Dorine Lampasona, Chloé Marillat, et Kim Nalin. Merci pour votre aide dans la construction des manips, et bravo de vous être accrochées malgré des paradigmes pas très évidents à comprendre.

Merci à la Team Souchay/Moulin et aux membres du méta-club avec qui il a été très agréable et enrichissant d'échanger sur la diversité de la mémoire et de la métacognition.

Parce que ma thèse n'aurait pas été la même sans vous, un immense merci à tous les doctorants du LPNC. Quelle chance incroyable d'avoir pu travailler avec des collègues aussi agréables et funs que vous. L'ambiance positive qui règne aura rendu ces 3 ans et demi tellement plus faciles. Vous êtes géniaux. Merci pour les moments partagés aux repas de midi, pour l'organisation des RJC, autour d'une bière, dans un match de foot, dans une soirée déguisée, pendant un jeu de carte (#teamcowboy), en conférence, ou autres.

Mention spéciale à Lucrèce, pour le partage de nos galères et nos doutes. Et Lise, la meilleure colloc' de colloque. Heureusement que j'ai pu compter sur toi pour m'accompagner sur des karaokés improvisés, sur une piste de danse, sur un bateau, ou les trois à la fois. J'espère à une prochaine, dans un sauna ou jacuzzi ?

Je remercie particulièrement mes collègues du E118. Quel bonheur d'avoir partagé ce bureau avec vous. Merci pour les échanges et conseils scientifiques mais aussi pour l'abondance de nourriture, les sessions potins, les débriefings sérieux, les rires, et pour l'ambiance générale qui donnait envie de venir travailler. Vous allez me manquer. Merci Audrey T., pour ta gentillesse, tu étais un peu la grande sœur du bureau, et merci pour tes recommandations sérieuses ; Martin, pour tes réflexions philosophiques et ton amour pour les muffins ; Thomas, le meilleur voisin, de m'avoir appris à jouer aux échecs, et pour ton énergie contagieuse ; Maryam, pour ton petit grain de folie, ta sensibilité et tes histoires toujours croustillantes ; Lucile, pour ton sens de l'orientation sur les pistes de ski et le partage des galères Meta-d'. Rémi, merci pour tes connaissances R, ton humour subtil, nos discussions sportives, mais surtout, les meilleurs gâteaux à la banane et au piment d'Espelette.

Un très grand merci aussi à la team des Docteurs, Brice, Elie, Olivier et Sam, pour les concours GeoGuessr, Wordle, Skribble qui auront animé mes premières années de thèse. Merci aussi pour les soirées Among Us interminables (une dernière ?), les après-midi jeux, les randos, les vacances, les Vendredi au bar, les matchs, etc. Merrick, parce qu'il n'y a que les idiots qui ne changent pas d'avis, j'ai eu de la chance de pouvoir apprendre à te connaître et de vivre les péripéties de cette thèse avec toi. Même nos disputes puériles me manqueront. Merci pour ton soutien pendant mes coups de stress.

Je remercie aussi Camille et Vaida. Je ne serais sûrement pas arrivée en thèse sans vous. Je suis si heureuse de vous avoir eu à mes côtés pendant ces années de Master. Malgré la distance, j'espère pouvoir continuer à vous voir vivre vos aventures, à devenir de grandes chercheuses, et à passer plein de bons moments avec vous.

Parce qu'une thèse n'est pas possible sans une aide extérieure, merci à tous mes amis pour leur soutien infaillible et tous les moments passés ensemble, le temps d'une heure, en week-end ou en vacances. Je suis tellement chanceuse de vous avoir dans ma vie. Je n'oublierai pas non plus votre aide dans la construction de mes expériences : merci d'avoir servi de cobayes pour le pilote de mes études, même quand elles n'étaient pas toujours très marrantes à réaliser. Je remercie plus particulièrement Nin', Mimi, Carousse et Teus qui auront suivi de près les différentes étapes de ma thèse et qui auront été mon meilleur réconfort dans les moments difficiles.

Enfin, le plus grand des remerciements à ma famille. Votre soutien est si important pour moi. Mille mercis à mes parents pour tous les sacrifices qu'ils ont faits pour que j'en arrive ici. Merci pour le temps passé à m'aider à réviser pendant mes études. Papa, merci de m'avoir soutenue et poussée à atteindre mes objectifs depuis toute petite. J'aurais tellement aimé que tu sois là pour vivre ça avec moi, mais je sais que tu aurais été fier. Maman et Auré, je ne sais pas ce que j'aurais fait sans vous. Merci d'être les femmes fortes que vous êtes, mes piliers. Je vous dois une partie de cette thèse, nos sessions de brainstorming pendant le confinement m'auront bien aidée à avancer. Merci aussi pour l'aide dans la relecture, et la préparation de mes présentations, mais plus particulièrement pour votre soutien au quotidien.

# Table des matières

Résumé.....	ii
Abstract .....	iii
Remerciements .....	iv
Liste des figures .....	ix
Liste des tableaux .....	x
Abréviations .....	xi
Préambule .....	xii
<b>PARTIE THÉORIQUE .....</b>	<b>2</b>
<b>Chapitre 1. Hétérogénéité de la métacognition .....</b>	<b>3</b>
1.1. Terminologie, définitions et modèles de la métacognition .....	3
1.1.1. Définitions et liens avec d'autres notions.....	3
1.1.2. Composants métacognitifs.....	4
1.1.3. Modèle de Nelson et Narens (1990).....	5
1.1.4. Mise en œuvre de la métacognition et utilité .....	7
1.2. Évaluation du monitoring.....	9
1.2.1. Paradigmes de monitoring .....	9
1.2.2. Mesures de monitoring .....	12
1.3. Conclusion du Chapitre 1 .....	15
<b>Chapitre 2. Bases des jugements métacognitifs.....</b>	<b>17</b>
2.1. Sources de l'auto-évaluation.....	17
2.1.1. Théorie de l'utilisation d'indices : présentation générale .....	17
2.1.2. Fluence de traitement : influence sur les JOL, les FOK, et les RCJ .....	19
2.1.3. Familiarité avec l'indice et récupération partielle de la cible : sources du FOK et des JOL différés.....	21
2.1.4. Remarques additionnelles et conclusions sur la théorie de l'utilisation d'indices.....	22
2.2. Généralité versus spécificité de la métacognition.....	24
2.2.1. Preuves comportementales .....	24
2.2.2. Preuves neuropsychologiques et neuro-fonctionnelles .....	26
2.3. Métacognition épisodique et sémantique : vision dichotomique.....	28
2.3.1. Mémoire épisodique et sémantique.....	28
2.3.2. Distinction épisodique et sémantique dans la métacognition .....	29
2.4. Conclusion du Chapitre 2.....	33
<b>Chapitre 3. Lien entre premier ordre et second ordre .....</b>	<b>35</b>
3.1. Lien statistique et conceptuel.....	35
3.1.1. Dépendance au niveau de la mesure ? .....	35
3.1.2. Dépendance au niveau cognitif ? .....	37
3.2. Traitement sémantique : utilisation dans la mémoire épisodique et la métacognition épisodique .....	41
3.2.1. Traitement sémantique : définition et modèles .....	41
3.2.2. Utilisation de la mémoire sémantique dans la mémoire épisodique .....	45
3.2.3. Utilisation de la mémoire sémantique dans la métacognition.....	50
3.3. Conclusion du Chapitre 3.....	54
<b>Problématique générale.....</b>	<b>56</b>

**PARTIE EXPÉRIMENTALE..... 58**

**Chapitre 4. Distinction entre feeling-of-knowing épisodique et sémantique dans le vieillissement..... 60**

4.1. Introduction.....	61
4.2. Methods.....	64
4.2.1. Systematic review and global meta-analysis .....	64
4.2.2. Analyses of individual scores from aggregated dataset .....	67
4.3. Results .....	69
4.3.1. Systematic review and global meta-analysis .....	69
4.3.2. Analyses of individual scores from aggregated dataset .....	71
4.4. Discussion du Chapitre 4 .....	74

**Chapitre 5. Étude de la métacognition dans la démence sémantique.....78**

5.1. Introduction.....	78
5.2. Méthode .....	84
5.2.1. Participants .....	84
5.2.2. Protocole .....	84
5.2.3. Matériel .....	88
5.2.4. Analyses.....	91
5.3. Résultats .....	91
5.3.1. Tâches de PoR .....	91
5.3.2. Tâches de RCJ .....	93
5.3.3. Questionnaire de plaintes mnésiques.....	95
5.4. Discussion du Chapitre 5 .....	96

**Chapitre 6. Importance du traitement sémantique pour les prédictions en mémoire épisodique .....100**

6.1. Introduction.....	101
6.2. Experiment 6.1: Predicting recognition of meaningless words .....	106
6.2.1. Method .....	107
6.2.2. Results .....	109
6.2.3. Interim discussion .....	112
6.3. Experiment 6.2: Improving processing of meaningless items.....	113
6.3.1. Method .....	113
6.3.2. Results .....	114
6.3.3. Interim Discussion.....	118
6.4. Experiment 6.3: Increasing familiarity with meaningless items.....	120
6.4.1. Method .....	121
6.4.2. Results .....	122
6.4.3. Interim discussion .....	125
6.5. Experiment 6.4: The role of recall .....	126
6.5.1. Method .....	127
6.5.2. Results .....	129
6.5.3. Interim discussion .....	132
6.6. General discussion of Chapter 6.....	133

<b>Chapitre 7. Effet de la fluence conceptuelle sur la métacognition épisodique.....</b>	<b>138</b>
7.1. Introduction.....	139
7.2. Expérience 7.1 : Effet de congruence sémantique sur les jugements rétrospectifs de confiance épisodique. ....	144
7.2.1. Méthode .....	144
7.2.2. Résultats .....	148
7.2.3. Discussion de l'expérience 7.1 .....	152
7.3. Expérience 7.2 : Effet de congruence sémantique sur les prédictions de reconnaissance en mémoire épisodique .....	155
7.3.1. Méthode .....	155
7.3.2. Résultats .....	159
7.3.3. Discussion de l'expérience 7.2. ....	161
7.4. Discussion générale du Chapitre 7 .....	162
<b>Chapitre 8. Discussion générale et perspectives.....</b>	<b>165</b>
8.1. Résumé des résultats .....	166
8.2. Apport des résultats aux théories de la métacognition .....	169
8.2.1. Lien entre déficit de premier ordre et sensibilité métacognitive.....	169
8.2.2. Influence des facteurs sémantiques dans la métacognition épisodique .....	170
8.3. Limites et perspectives.....	173
<b>Conclusion de la thèse .....</b>	<b>178</b>
<b>Annexes.....</b>	<b>179</b>
Annexe A. Chapitre 4 : Résumé des études incluses dans la méta-analyse globale. ....	179
Annexe B. Chapitre 5 : Liste des mots utilisé pour la tâche de PoR sémantique. ....	180
Annexe C. Chapitre 5 : Tableau récapitulatif des moyennes (écart-types) pour chaque tâche et groupe. ....	181
Annexe D. Chapitre 6 : Exemples de paires utilisées pour un bloc.....	182
Annexe E. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.1. ....	183
Annexe F. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.2. ....	184
Annexe G. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.3. ....	185
Annexe H. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.3. ....	186
Annexe I. Chapitre 6 : Résultats des analyses sur l'efficacité métacognitive. ....	187
Annexe J. Chapitre 7 : Tableaux récapitulatif des résultats de l'expérience 7.1. ....	192
<b>Bibliographie.....</b>	<b>193</b>

# Liste des figures

Figure 1.1 : Modèle de Nelson & Narens (1990).	6
Figure 1.2 : Exemples de jugements métacognitifs étudiés sur des tâches de mémoire (liste non exhaustive).	12
Figure 1.3 : Logique de la TDS.	14
Figure 2.1 : Types d'indices à la base des jugements métacognitifs selon Koriat (1997).	19
Figure 3.1 : Exemples d'interprétations sur les différences entre performance de premier ordre et sensibilité métacognitive (figure issue de Rouault et al., 2018).	36
Figure 3.2 : Exemples de modèles cognitivistes.	42
Figure 3.3 : Théories neuro-inspirées de l'organisation du système sémantique. Schémas issus de Patterson et al. (2007).	44
Figure 3.4 : Modèle SPI (issu de Tulving, 1995).	46
Figure 3.5 : Paradigme utilisé dans la thèse de Mazancieux (2020).	52
Figure 4.1: Flow chart of the steps for identification and selection of records included in the systematic review and meta-analysis.	66
Figure 4.2 : Forrest plot of the effect of eFOK deficit in OA.	70
Figure 4.3 : Forrest plot showing an absence of sFOK deficit in OA.	71
Figure 4.4 : Funnel plot centered on the overall effect size (vertical line) for eFOK model (a) and sFOK model (b).	71
Figure 4.5 : Forrest plots of the different effects analyzed from aggregated dataset.	72
Figure 4.6 : Forrest plots of the effect of eFOK deficit in OA using other measures.	73
Figure 4.7 : Forrest plots of of eFOK deficit in OA using median-splits based on recognition.	74
Figure 5.1 : Modèle CAM, reproduit d'après Morris & Hannesdottir (2004).	80
Figure 5.2 : Procédures utilisées pour les tâches de PoR.	87
Figure 5.3 : Procédures utilisées pour les tâches de RCJ.	88
Figure 5.4 : Résultats sur la tâche de PoR.	93
Figure 5.5 : Résultats sur la tâche de RCJ.	95
Figure 6.1 : Procedure of one block of Experiment 6.1.	108
Figure 6.2 : Results of Experiment 6.1.	110
Figure 6.3 : Results of Experiment 6.2.	115
Figure 6.4 : Results of Experiment 6.3.	124
Figure 6.5 : Results of Experiment 6.4.	130
Figure 7.1 : Procédure utilisée dans l'Expérience 7.1.	146
Figure 7.2 : Design et exemples d'images utilisées dans l'expérience 7.1.	147
Figure 7.3 : Résultats de l'expérience 7.1.	150
Figure 7.4 : Distributions postérieures des effets moyens ( $\mu$ , ligne en pointillés) de la congruence de la cible (A), de la congruence du distracteur (B), et de leur interaction (C).	151
Figure 7.5 : Procédure utilisée dans l'Expérience 7.2.	157
Figure 7.6 : Résultats de l'expérience 7.2.	160
Figure 8.1 : Proposition d'un modèle théorique récapitulant les résultats de la littérature contribuant à la compréhension des relations entre processus de premier ordre (en bas) et de second ordre (en-haut) selon que les jugements métacognitifs soient prédictifs (à gauche) ou rétrospectifs (à droite) (schéma inspiré de Mazancieux et al., soumis).	166
Figure 8.2 : Intégration des résultats des 2 méta-analyses du Chapitre 4.	169
Figure 8.3 : Intégration des résultats du Chapitre 5.	170
Figure 8.4 : Intégration des résultats des Chapitres 6 (a) et 7 (b et c).	172

# Liste des tableaux

---

Tableau 1.1 : Mesures de monitoring décrites dans la littérature.....	15
Tableau 4.1 : Available data for each individual of the aggregated dataset.....	68
Tableau 5.1 : Données démographiques des participants du projet META-DEM.....	85
Tableau 5.2 : Descriptif des personnes DS : genre, âge, latéralité de l'atteinte, et observations particulières pendant la passation. ....	86
Tableau 5.3 : Récapitulatif des tâches de métacognition et du matériel utilisé.....	90
Tableau 6.1 : Summary of the 4 experiments of Chapter 6.....	106
Tableau 6.2 : Summary of results from Chapter 6 .....	133
Tableau 7.1 : Récapitulatif des deux expériences.....	143
Tableau 8.1 : Récapitulatif des résultats de cette thèse.....	168

# Abréviations

---

<b>AUROC2</b>	Type-2 area under ROC curve (Aire sous la courbe ROC de type-2)
<b>BREF</b>	Batterie rapide d'évaluation frontale
<b>CAM</b>	Cognitive Awareness Model
<b>DS</b>	Démence sémantique
<b>DFT</b>	Démence fronto-temporale
<b>EOL</b>	Ease of learning (jugement de facilité d'apprentissage)
<b>eFOK</b>	Feeling of knowing épisodique
<b>eRCJ</b>	Jugement rétrospectif de confiance épisodique
<b>ePoR</b>	PoR épisodique
<b>FOK</b>	Feeling of knowing
<b>G</b>	Guess
<b>HDI</b>	Highest Density Interval
<b>IRMf</b>	Imagerie par résonance magnétique fonctionnelle
<b>JOL</b>	Judgment of learning (jugement d'apprentissage)
<b>K</b>	Know
<b>MMSE</b>	Mini-Mental State Examination
<b>MA</b>	Maladie d'Alzheimer
<b>ME</b>	Mémoire épisodique
<b>MS</b>	Mémoire sémantique
<b>OA</b>	Older adults
<b>PoR</b>	Prediction of recognition (prédiction de reconnaissance)
<b>PW</b>	Pseudoword
<b>R</b>	Remember
<b>RCJ</b>	Retrospective confidence judgment; jugement de confiance rétrospectif
<b>sFOK</b>	Feeling of knowing sémantique
<b>sRCJ</b>	Jugement rétrospectif de confiance sémantique
<b>sPoR</b>	PoR sémantique
<b>TDS (ou SDT)</b>	Théorie de détection du signal
<b>vcDFT</b>	Variante comportementale de la DFT
<b>W</b>	Word
<b>YA</b>	Young adults

# Préambule

---

La capacité des individus à évaluer et contrôler leur propre cognition (e.g., perception, action, mémoire) renvoie au concept de métacognition. Dans cette thèse, l'intérêt était porté sur les capacités de monitoring, c'est-à-dire d'auto-évaluation, sur des tâches de mémoire. Ce projet a été construit sur l'observation que, dans la littérature, des travaux mettent en évidence une distinction entre les capacités de monitoring selon que la tâche cognitive porte sur la mémoire épisodique (i.e., souvenirs contextualisés dans le temps et l'espace) ou la mémoire sémantique (i.e., connaissances générales). Cet effet a notamment été montré dans des populations cliniques comme la maladie d'Alzheimer ou dans le vieillissement sain. Cependant, toutes les études ne mènent pas aux mêmes conclusions. Pour cela, dans le Chapitre 4, nous proposons une brève revue de la littérature sur cette question dans le vieillissement sain.

Plusieurs théories ont été proposées pour expliquer ce décalage des performances métacognitives entre tâches épisodiques et sémantiques. Notamment, certaines de ces théories postulent que les processus de mémoire mis en jeu dans la tâche mnésique interviennent également dans les jugements métacognitifs (e.g., recollection). Par conséquent, un déclin de ces processus pourrait également entraîner un déficit au niveau métacognitif. En parallèle, dans le champ de la métacognition au-delà de la métamémoire, la relation entre performance sur la tâche cognitive (i.e., de premier ordre) et capacités métacognitives (i.e., de second ordre) est un débat d'actualité. Ainsi, l'objectif de cette thèse était d'apporter des éléments supplémentaires pour comprendre comment la mémoire à long terme intervient et est utilisée pour former des jugements de monitoring portés dans des tâches de mémoire. Nous cherchions à comprendre si un déclin d'une forme de mémoire (i.e., épisodique ou sémantique) était à l'origine d'une diminution des capacités de monitoring sur des tâches portant sur ce type de mémoire (Objectif 1).

De nombreux travaux ont également remis en question une distinction stricte entre mémoire épisodique et sémantique. En plus de partager des points communs au niveau neuronal et fonctionnel, les deux formes de mémoire interagissent. Plus particulièrement, les traitements sémantiques sont mis en jeu dans la création et la récupération de souvenirs et donc dans la mémoire épisodique. De plus, quelques travaux ont mis en évidence l'influence des connaissances sémantiques sur certains jugements métacognitifs. Cependant, la majorité de ces études ont été menées sur des jugements d'apprentissage (JOL) ou avec du matériel verbal. Ainsi, plus d'études sont nécessaires afin de mieux comprendre l'étendue de cette influence, avec par exemple du matériel visuel, mais également afin de voir si le traitement sémantique des informations est nécessaire pour pouvoir évaluer et prédire précisément sa propre capacité à récupérer des souvenirs en mémoire. Ainsi, un second objectif de ce projet était de mieux comprendre comment les connaissances sémantiques peuvent être utilisées pour juger ses propres performances en mémoire, y compris dans des tâches épisodiques (Objectif 2).

Pour simplifier les formulations, nous appellerons jugements métacognitifs (ou le type de jugement plus précis ; e.g., *feeling-of-knowing*) épisodiques/sémantiques les jugements métacognitifs faits sur des tâches de premier ordre portant sur la mémoire d'informations nouvelles (i.e., épisodique) ou de connaissances préexistantes (i.e., sémantique). Ainsi, nous ne considérons pas nécessairement que les jugements faits sur une tâche d'une forme de mémoire (i.e., épisodique ou sémantique) fassent exclusivement appel à des processus ou des informations de ce type de mémoire.

Ce manuscrit de thèse se décompose en 8 chapitres. Les trois premiers chapitres constituent le cadre théorique et permettent d'introduire les chapitres expérimentaux. Le Chapitre 1 vise à définir plus précisément ce qu'est la métacognition, et en particulier le monitoring (objet de cette thèse), mais aussi à présenter les différents paradigmes généralement utilisés pour l'évaluer. Ce chapitre présente également un détail des mesures de métacognition, dans le but d'introduire celles qui ont été utilisées dans les expériences de cette thèse, mais également pour aider à la compréhension des études présentées et décrites tout au long du manuscrit. Le Chapitre 2 présente les bases des jugements métacognitifs, les différents indices les influençant ainsi que les résultats argumentant une possible dissociation entre métacognition épisodique et sémantique. Enfin, le Chapitre 3 introduit l'idée d'une interaction entre mémoires épisodique et sémantique, et présente plus précisément une description du système de traitement sémantique. Dans ce chapitre, nous verrons également comment le traitement sémantique des informations peut influencer à la fois la mémoire épisodique et les jugements métacognitifs épisodiques. Nous terminerons cette partie théorique par une présentation de la problématique générale de la thèse.

La section expérimentale comprend les Chapitres 4, 5, 6 et 7. Dans le Chapitre 4, nous présentons deux méta-analyses sur le *feeling-of-knowing* dans le vieillissement. Le but était de rendre compte des performances métacognitives des personnes âgées sur des tâches épisodiques et sémantiques. La première méta-analyse porte sur toutes les études publiées sur le sujet dans la littérature. La seconde inclut uniquement des données publiées et non publiées issues de notre laboratoire, et visait à évaluer plus précisément le rôle des capacités en rappel et en reconnaissance dans les effets observés, à l'aide de données agrégées par participant. Ce chapitre visait à répondre plus particulièrement à l'Objectif 1 de cette thèse.

Le Chapitre 5 présente une étude multicentrique visant à étudier les performances métacognitives de personnes atteintes de démence sémantique, une pathologie neurodégénérative altérant principalement les connaissances sémantiques. Cette pathologie étant rare et afin de pouvoir recruter un plus grand nombre de participants, ce projet a été réalisé en collaboration avec 3 centres hospitaliers français : Grenoble, Rennes et Angers. Pour plus d'homogénéité, j'ai effectué les déplacements dans ces centres afin de récolter les données auprès des patients. Ainsi, ce projet a été monté en prenant en compte les contraintes liées à la distance entre ces centres qui limitaient le nombre de déplacements possibles (e.g., nombre d'items par tâches pour éviter des temps de passation trop longs qui auraient requis plusieurs sessions de passation par patient). Les données de ces patients ont été comparées à celles de personnes âgées saines et de quelques personnes atteintes de maladie d'Alzheimer, recrutées sur Grenoble. Ce Chapitre 6 visait à répondre plus particulièrement à l'Objectif 1 de cette thèse, mais apporte des informations pour l'Objectif 2 également.

Le Chapitre 6 regroupe 4 expériences (6.1, 6.2, 6.3, et 6.4) visant à investiguer la nécessité des traitements sémantiques pour prédire ses performances en mémoire. Le même paradigme a été utilisé dans ces 4 expériences, avec quelques modifications d'une expérience à l'autre. Ce chapitre concernait l'Objectif 2 de cette thèse.

Enfin, le Chapitre 7 présente deux expériences (7.1 et 7.2) ayant pour but d'approfondir comment les individus utilisent leur connaissance du monde pour évaluer leur mémoire épisodique. Deux paradigmes ont été utilisés : un paradigme de jugements rétrospectifs (Expérience 7.1) et un de prédictions (Expérience 7.2). Pour les besoins de ces deux expériences, une base d'images a été créée (voir [Box 2](#) et [3](#)). Ce Chapitre 7 répond à l'Objectif 2 de cette thèse.

Le Chapitre 8 est consacré à la Discussion Générale de la thèse. Nous proposons un résumé de l'ensemble des résultats obtenus dans les différentes études réalisées ainsi que des perspectives.

Le Chapitre 4 est un article ayant été soumis à la revue *Scientific Reports* et est en cours de review. Le Chapitre 6 est un article en cours d'écriture et sera bientôt soumis à la revue *Journal of Experimental Psychology: Learning & Memory*. Ces chapitres ayant donc été travaillés dans leur version anglaise, ils ont été gardés en **anglais** et reformattés pour les besoins de ce manuscrit de thèse. Si besoin, une version traduite en français pourra être fournie. Le reste du manuscrit, y compris les Chapitres 5 et 7, a été rédigé en **français**.

Des conclusions ou résumés intermédiaires, ou des notes sont présentés dans des encarts gris. Les conclusions générales des chapitres sont présentées dans des encarts bleus.

Cette thèse a été financée par la bourse MESRI de l'EDISCE de l'Université Grenoble Alpes et a été conduite au sein du Laboratoire de Psychologie et de NeuroCognition de Grenoble.

**A**vez-vous déjà discuté avec vos proches et douté d'un souvenir ou d'une connaissance partagée ? Imaginez que quelques années après avoir voyagé à Florence, vous racontez votre visite du Musée des Offices à des amis. Lorsque vous indiquez avoir trébuché devant la *Joconde*, vos proches s'étonnent. Pour vous défendre, vous vous justifiez : « Je reconnais que c'est étrange, et je suis en fait sûre à 50% que c'était bien ce tableau-là. », alors qu'il s'agissait en réalité bien de ce tableau, en exposition temporaire\*. En continuant le résumé de votre visite, vous expliquez avoir été surprise par la beauté d'un célèbre tableau, autour duquel de nombreuses personnes étaient réunies, mais au moment de le décrire vous ne retrouvez plus duquel il s'agissait. Vous avez alors la forte sensation de pouvoir vous en souvenir, un *feeling-of-knowing*, et quelques informations vous reviennent : il a été peint par Sandro Botticelli et son nom comporte le nom d'une déesse. Après des propositions et descriptions données par vos proches, vous arrivez à vous rappeler qu'il s'agissait de la *Naissance de Vénus* et qu'il y avait une femme aux cheveux longs debout sur un coquillage.

Dans cet exemple, plusieurs jugements ont été portés sur vos propres souvenirs épisodiques (e.g., degré de confiance en son exactitude, sensation de pouvoir vous souvenir). Ces auto-évaluations et sensations peuvent s'assimiler à des manifestations naturelles du fonctionnement métacognitif et plus particulièrement du monitoring. Dans cette thèse, le but était de mieux comprendre comment nous arrivons à auto-évaluer la validité de nos souvenirs et pourquoi nous sommes parfois amenés à faire des erreurs dans cette évaluation. Nous étudierons comment notre mémoire est utilisée pour auto-évaluer l'exactitude de nos souvenirs. Par exemple, est-ce que la connaissance que la *Joconde* se trouve généralement au Louvre de Paris a influencé la confiance que vous avez eu en l'exactitude de votre souvenir ? Auriez-vous eu plus confiance en ce souvenir si vous aviez trébuché au Louvre, un contexte plus congruent pour la *Joconde* ? Dans ce travail, nous avons notamment cherché à comprendre comment les connaissances que nous avons sur le monde qui nous entoure (i.e., mémoire sémantique) nous influence dans notre façon d'évaluer l'exactitude de nos souvenirs. De plus, auriez-vous eu la même sensation de pouvoir vous souvenir du tableau si celui-ci n'était pas aussi connu et si vous n'en aviez jamais entendu parler avant de le voir ? Ainsi, nous avons également cherché à voir si ces informations sémantiques à propos de nos souvenirs (e.g., nom du peintre) sont nécessaires afin de prédire correctement si nous serons capables de les récupérer dans un futur proche, avec de l'aide.

\* Histoire inventée, bien que La Joconde aurait effectivement été exposée une fois au musée des Offices, entre 1911 et 1914, à la suite d'un vol. (source : <https://www.beauxarts.com/grand-format/ep-1-le-jour-ou-la-joconde-sest-fait-la-belle/>)

# PARTIE THÉORIQUE

# Chapitre 1.

## Hétérogénéité de la métacognition

---

---

### 1.1. Terminologie, définitions et modèles de la métacognition

#### 1.1.1. Définitions et liens avec d'autres notions

Bien qu'un intérêt ait été porté sur la capacité à réfléchir sur soi et sur ses propres expériences depuis très longtemps (e.g., devise "connais-toi toi-même" de Socrate), la métacognition a été plus clairement définie par Flavell (1976). Ce terme a été proposé pour renvoyer à la connaissance et à l'autorégulation de nos propres activités cognitives et apprentissages (Flavell, 1976, 1979). Il s'agit d'une compétence de haut-niveau permettant de « penser sur la pensée » (i.e., *thinking about thinking* en anglais) et de comprendre comment le système cognitif fonctionne et opère. La métacognition peut être étudiée dans de nombreux domaines de la cognition comme la perception, les apprentissages, la mémoire, etc. Ce projet de thèse a pour objectif d'étudier les capacités métacognitives mises en jeu lors de la réalisation de tâches de mémoire, ce qui renvoie à la métamémoire (Flavell, 1971), un terme parfois utilisé pour renvoyer à la connaissance et la régulation de nos propres processus mnésiques. Dans les expériences de cette thèse, nous utiliserons le terme de métacognition pour renvoyer à la métacognition de la mémoire.

La métacognition est liée à d'autres concepts. Notamment, l'idée de processus de haut-niveau qui nécessitent une représentation d'un état de plus bas niveau est présente dans certaines définitions de la métacognition et est également utilisée pour l'étude de la conscience par certains auteurs (e.g., Rosenthal, 2004). Selon ces auteurs, l'*awareness*<sup>1</sup> consciente est la perception du monde qui s'accompagne d'une expérience subjective (Lau & Rosenthal, 2011). La recherche sur la métacognition fournit un cadre pour investiguer la conscience des états mentaux (Rosenthal, 2000) et mieux comprendre les processus sous-tendant la validité des expériences subjectives, bien que les phénomènes métacognitifs ne soient pas toujours accompagnés d'une expérience consciente (Kentridge & Heywood, 2000). Plus particulièrement, l'évaluation du monitoring et du contrôle, deux processus métacognitifs (Nelson & Narens, 1990, voir Figure 1.1 plus loin) permettent d'appréhender plus précisément les phénomènes subjectifs ressentis à propos d'une activité cognitive (Koriat, 2007; Nelson, 1996). Dans un contexte plus clinique, l'étude de la métacognition peut également servir d'outil

---

<sup>1</sup> Le terme anglais est ici utilisé car les termes *consciousness* et *awareness* ont la même traduction française (i.e., conscience).

pour approfondir la compréhension de l'anosognosie de manière expérimentale et inversement (Ernst et al., 2015; Sunderaraman & Cosentino, 2017). L'anosognosie est un symptôme se définissant par l'absence de conscience d'un trouble (Babinski, 1914). Initialement proposé pour renvoyer à l'inconscience d'une faiblesse ou paralysie d'un membre, ce terme a été étendu aux troubles cognitifs. Notamment, on observe dans les stades avancés de certaines maladies neurodégénératives (e.g., maladie d'Alzheimer, MA) une diminution de la conscience de soi et notamment de conscience de leurs troubles (e.g., Alzheimer, 1907; voir Ernst et al., 2015 pour une revue dans la MA). Bien que leur étude se soit développée en parallèle, l'anosognosie (clinique) et la métacognition (psychologie cognitive) renvoient à la connaissance et à la conscience de soi. L'étude de la métacognition permet de comprendre comment le soi opère et comment un individu réfléchit sur ses performances. Ainsi, les paradigmes de métacognition fournissent un cadre permettant d'étudier l'anosognosie dans des pathologies de la cognition (voir [Chapitre 5](#)).

### **1.1.2. Composants métacognitifs**

Comme introduit précédemment, un aspect fondamental dans l'étude de la métacognition est son hétérogénéité. Plusieurs types de composants métacognitifs ont été décrits dans la littérature (Baker, 1991; A. Brown, 1987; Flavell, 1979; Schraw & Moshman, 1995; voir S. Chen & McDunn, 2022 et Lai, 2011 pour des revues) telles que les connaissances et la régulation métacognitives (Schraw & Moshman, 1995). D'un côté, les connaissances métacognitives font référence à toutes les croyances (correctes ou erronées) implicites ou explicites que les personnes possèdent sur le fonctionnement cognitif et sur ce qui l'influence, au niveau intra-individuel, interindividuel ou universel (Flavell, 1979). Par exemple, un individu peut estimer qu'il mémorise mieux les dates de faits historiques que les dates d'anniversaires de ses proches (croyance intra-individuelle), ou encore que son ami est meilleur que lui pour retenir le nom des endroits visités pendant leurs vacances (croyance interindividuelle). Il peut également savoir que les mots concrets sont généralement mieux retenus que les mots abstraits (croyance universelle). Ces croyances peuvent porter sur les individus, la tâche réalisée ou les stratégies utilisées. D'autres classifications de ces croyances ont été proposées, comme Schraw & Moshman (1995) qui les classent selon leur fonctionnalité : déclaratives (ce qui influence les performances), procédurales (comment réguler la cognition et utiliser les stratégies) et conditionnelles (quand et pourquoi utiliser les stratégies).

D'un autre côté, la régulation de la cognition regroupe le monitoring des processus cognitifs, leur contrôle et la réflexion portée sur ces processus (Schraw, 1998; Schraw & Moshman, 1995). Ce composant inclut la planification des activités cognitives, comme l'adaptation et la mise en place de stratégies d'apprentissage adaptées, et l'allocation des

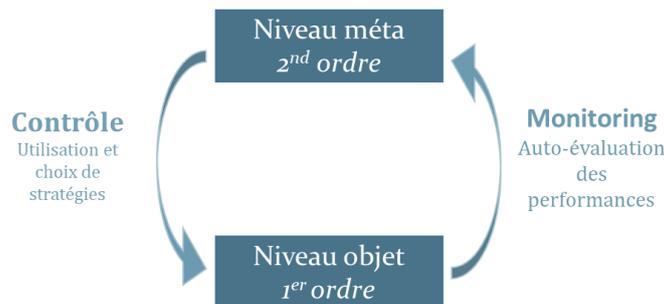
ressources temporelles et cognitives dans une tâche. Un autre aspect de la régulation est l'évaluation, ou monitoring (Nelson & Narens, 1990), c'est-à-dire l'appréciation des processus cognitifs et l'estimation des performances et des résultats d'un apprentissage, et qui sera l'objet de cette thèse. Le monitoring peut également se manifester par des expériences métacognitives, qui renvoient aux sensations subjectives ressenties vis-à-vis d'un phénomène cognitif (Flavell, 1979). Ces sensations informent l'individu sur le résultat ou l'état d'une activité cognitive. Un exemple est la sensation du mot sur le bout de la langue (A. S. Brown, 1991; R. Brown & McNeill, 1966; Moulin & Souchay, 2014). Cette expérience se manifeste par l'incapacité objective de rappeler un mot, associée à la sensation subjective de connaître un mot et d'être sur le point de le retrouver (Koriat, 2000). D'autres exemples d'expériences métacognitives sont les sensations de familiarité (Jacoby & Dallas, 1981; Mandler, 1980), de savoir (*feeling-of-knowing*, FOK; Hart, 1965; Koriat et al., 2004; Read & Bruce, 1982; Reder, 1988), d'oubli (Halamish et al., 2011), de certitude (W. F. Brewer & Sampaio, 2012; Koriat, 2008), ou encore la fluence de traitement (Whittlesea, 1993).

Les différents composants de la métacognition décrits interagissent et s'influencent l'un l'autre. Notamment, les expériences interviendraient dans la construction des connaissances métacognitives et ces connaissances guideraient aussi la régulation cognitive. Ces phénomènes aident à optimiser nos ressources cognitives car ils permettent de traiter ce que le système cognitif a appris, l'importance de ce qui est appris, ce qu'il a retenu, et permet de juger et de discriminer ce qui doit être retenu de ce qui peut être oublié (Kralik et al., 2018). Ils permettent également d'agir et d'adapter les stratégies pour répondre à nos objectifs et à nos buts (Flavell, 1979). Cependant, les croyances et les expériences métacognitives peuvent être erronées et amener un sujet à se tromper ou à ne pas utiliser les stratégies les plus efficaces. Par exemple, un étudiant peut lire plusieurs fois un cours et estimer que c'était la meilleure stratégie pour bien le retenir puis se retrouver en partiel et se rendre compte qu'il a oublié la moitié de ce qu'il a lu, alors qu'il aurait pu être plus efficace qu'il se teste avant et qu'il essaye de rappeler son cours alors qu'il était encore chez lui pour voir ce qu'il oubliait. Ainsi, des connaissances ou une régulation erronées peuvent entraîner une mauvaise allocation de l'effort cognitif ou des ressources temporelles.

### **1.1.3. Modèle de Nelson et Narens (1990)**

L'hétérogénéité de la métacognition est également illustrée dans les modèles qui tentent de la décrire. Le modèle conceptuel de métacognition le plus classiquement utilisé (Nelson & Narens, 1990, voir Figure 1.1) suggère que les activités cognitives peuvent être divisées en deux niveaux qui interagissent : le niveau objet et le niveau méta. Les souvenirs et les

informations à apprendre ou ayant été apprises, résultant d'un traitement cognitif, constituent le niveau objet. D'un autre côté, la représentation subjective de ces éléments et de leur exactitude se situent au niveau méta. Ce niveau méta contient un modèle (i.e., représentation mentale) de ce qui fait partie du niveau objet. Le modèle suppose également que les deux niveaux communiquent par deux processus distincts de régulation : le monitoring et le contrôle. Ces deux processus diffèrent dans la direction du flux d'informations qu'ils contiennent. Le processus de monitoring sert à informer le niveau méta de l'état du niveau objet (flux du niveau objet au niveau méta). Il permet ainsi de superviser ce qui se situe et ce qui est réalisé au niveau objet. Ainsi, le monitoring permet à une personne d'évaluer l'exactitude de ses souvenirs et des informations qu'elle récupère en mémoire. En complément, le processus de contrôle coordonne et influence le niveau objet ou son état, en initiant, ajustant ou terminant une action (flux du niveau méta au niveau objet). Plus concrètement, dans le cadre d'un apprentissage, ou plus largement d'une activité cognitive, ce processus de contrôle se met en œuvre par la mise en place ou l'ajustement de stratégies afin de modifier l'état du niveau objet. Ce contrôle permet d'ajuster le niveau objet en vue de répondre à un objectif et peut dépendre de l'évaluation faite par le processus de monitoring (Efklides, 2006). Par exemple, en préparant son examen du lendemain, un étudiant peut faire le point sur ce qu'il a retenu du cours qu'il révise (i.e., monitoring). Sur la base de cette auto-évaluation, l'étudiant peut ensuite décider de repasser du temps à apprendre la partie qu'il juge comme moins maîtrisée (i.e., contrôle).



**Figure 1.1** : Modèle de Nelson & Narens (1990).  
Les flèches reflètent l'information qui circule entre les deux niveaux.

Les processus de monitoring et de contrôle du modèle de Nelson & Narens (1990) sont des exemples de régulation métacognitive, mais ils peuvent également être influencés par les connaissances métacognitives. Par exemple, ces connaissances peuvent influencer des jugements de monitoring (e.g., « cet évènement a eu lieu mercredi dernier, mais je ne suis pas certaine car j'ai une très mauvaise mémoire des dates ») et le contrôle (e.g., « je vais vérifier mon agenda pour savoir si je suis disponible demain soir car je ne suis pas forte pour retenir

mon emploi du temps »). De plus, la limite entre les processus qui relèvent du niveau cognitif (i.e., objet) et ceux qui relèvent du niveau méta est parfois floue puisque certains mécanismes sont mis en jeu dans les deux niveaux. Par exemple, le processus de recollection (i.e., capacité à se remémorer des détails spécifiques d'un événement vécu) qui est au cœur de la mémoire épisodique (ME) peut à la fois servir au niveau objet car il aide à se souvenir d'un épisode, mais il peut aussi servir de source aux jugements de monitoring. Ce point sera développé dans le [Chapitre 3](#) lorsque nous aborderons les relations entre mémoire et métacognition.

La description des processus de monitoring et de contrôle a été reprise dans des modèles plus récents de la métacognition, dans lesquels d'autres processus métacognitifs, tels que l'évaluation, la planification et la simulation mentale, ont été décrits (Kralik et al., 2018). Dans cette thèse, nous nous focalisons sur le processus de monitoring, c'est-à-dire l'évaluation des performances sur une tâche de mémoire.

### **1.1.4. Mise en œuvre de la métacognition et utilité**

Bien que nous réfléchissions parfois intentionnellement à nos performances et nos processus cognitifs, nos capacités métacognitives sont la plupart du temps mises en jeu de manière involontaire et automatique (Martinez, 2006). Par exemple, le mot-sur-le-bout-de-la-langue qui est considéré par certains (e.g., Schwartz, 1999, 2006; Schwartz & Frazier, 2005) comme une expérience métacognitive, arrive automatiquement. Ces capacités interviennent dans de nombreux aspects de la vie de tous les jours. Leur importance a été largement étudiée dans le cadre de l'éducation. Par exemple, plusieurs études ont montré un rôle bénéfique de l'utilisation de la métacognition et de l'apprentissage auto-régulé pour l'amélioration des performances scolaires (e.g., Dignath & Büttner, 2008). La métacognition permet de réguler les apprentissages (Bjork et al., 2013). La recherche portant sur la répartition du temps d'étude met en avant le rôle de la métacognition dans les choix de gestion du temps, de sélection des informations et de leur ordre lors d'un apprentissage (voir Son & Kornell, 2008 pour une revue). Par exemple, quand les participants avaient la possibilité de réétudier des paires d'items déjà apprises, ils avaient plus tendance à choisir les items qu'ils étaient incapables de rappeler avant. Ils passaient également plus de temps à réapprendre les items jugés comme étant difficiles que faciles, ou qu'ils étaient moins confiants de rappeler, bien que ce choix dépendaient des conditions de passation et des objectifs fixés en amont (e.g., Hanczakowski et al., 2014; Metcalfe & Finn, 2008b; Son & Metcalfe, 2000). Ces études mettent en évidence l'importance des capacités de monitoring et de contrôle dans les apprentissages, puisque le monitoring (e.g., estimation de la difficulté ou prédiction d'une récupération en mémoire future dans ces études) influence le contrôle (e.g., stratégies d'apprentissage) qui peut permettre d'optimiser et améliorer les apprentissages.

Le rôle de la métacognition ne se limite pas aux apprentissages (voir Heyes et al., 2020). Des travaux suggèrent notamment une importance des jugements métacognitifs dans les témoignages oculaires (e.g., Berkowitz et al., 2022; Deffenbacher, 1980). En effet, les décisions des juges lors de procès sont influencées par la certitude que les témoins ont de leur identification des coupables, bien qu'un témoin qui identifie une personne avec certitude ne garantisse pas que cette identification soit correcte (Sauer et al., 2019). La métacognition semble également jouer un rôle dans le déchargement cognitif (*cognitive offloading* en anglais, Hu et al., 2019), c'est-à-dire les choix de reposer sur des ressources de l'environnement extérieur pour réduire la demande cognitive (e.g., noter ses rendez-vous dans un agenda). Par exemple, les participants qui avaient moins confiance en leur capacité future de rappel avaient plus tendance à reposer sur une aide extérieure (ici, sauvegarder les paires de mots à apprendre pour le test; Hu et al., 2019). Le rôle de la métacognition a également été étudié dans d'autres domaines (voir Norman et al., 2019 pour une revue) comme les interactions sociales (e.g., Frith, 2012) et certaines pathologies (e.g., Seow et al., 2021; Wells, 2011).

Notons cependant que la métacognition n'est pas toujours bénéfique puisque certains auteurs soulignent des limites voire des aspects néfastes à l'étude de la métacognition (Norman, 2020). Notamment, cette auteure propose trois façons dont la métacognition pourrait ne pas toujours être bénéfique : 1) verbaliser ses expériences métacognitives lors de la réalisation d'une tâche cognitive pourrait interférer avec les performances sur cette tâche ; 2) le coût de la mise en place de stratégies métacognitives peut surpasser ses bénéfices ; et 3) les évaluations métacognitives négatives (e.g., penser être moins compétent qu'une autre personne) peuvent diminuer le sentiment d'auto-efficacité et avoir des conséquences sur l'estime de soi.

En résumé, l'évaluation et la gestion de nos propres activités cognitives guident nos comportements et peuvent avoir des conséquences sur plusieurs aspects de notre vie. Cependant, nous sommes parfois amenés à faire des erreurs d'évaluation ou à ne pas adopter les meilleurs comportements en réponse à ces évaluations (Bjork et al., 2013). Selon Koriat (2007), la recherche expérimentale sur la métacognition a essayé de répondre à cinq questions fondamentales :

- 1) Quelles sont les bases des jugements métacognitifs qui évaluent les souvenirs ?
- 2) À quel point ces jugements sont-ils corrects ?
- 3) Quels processus sous-tendent l'exactitude de ces jugements ?
- 4) Comment le monitoring influence-t-il le contrôle ?
- 5) Comment les processus métacognitifs influencent-ils les performances réelles ?

Cette thèse s'inscrit dans la lignée des travaux essayant de répondre aux trois premières questions.

## 1.2. Évaluation du monitoring

### 1.2.1. Paradigmes de monitoring

De nombreuses recherches ont été conduites afin de mieux comprendre comment les personnes arrivent à réfléchir sur leur activité mnésique et leurs performances. Des méthodes de mesure indirectes de monitoring sont parfois utilisées (e.g., mouvements oculaires), notamment dans des populations ne pouvant pas verbaliser leur réponse comme les enfants, les personnes présentant un trouble cognitif, ou encore les animaux (e.g., Kiani & Shadlen, 2009; Roderer & Roebbers, 2010). Dans cette thèse, nous nous sommes intéressés aux jugements explicites de monitoring afin de comprendre leurs bases. Plusieurs paradigmes ont été développés pour mesurer de manière explicite et contrôlée le processus de monitoring *online* (i.e., pendant la réalisation d'une tâche de mémoire, et non par un questionnaire portant sur la vie quotidienne). La méthode classiquement utilisée consiste à demander aux personnes de faire un jugement explicite sur leur performance en mémoire (Hart, 1965; Nelson & Narens, 1990). Dans ces paradigmes, les participants sont amenés à donner leur avis subjectif sur leur réussite ou échec à la tâche de mémoire. Ainsi, il est possible de mesurer à la fois la performance mnésique (de premier ordre) et les performances métacognitives (de second ordre). La performance de premier ordre est évaluée sur la base des réponses dites de type-1 (i.e., réponses au test de mémoire), qui reposent sur les processus cognitifs du niveau objet du modèle de Nelson & Narens (1990, voir Figure 1.1). D'un autre côté, la performance de second ordre reflète le monitoring et est évaluée en comparant la performance mnésique (i.e., de premier ordre) aux jugements de type-2 portés sur les réponses de type-1. Dans une tâche de mémoire classique, la réponse de type-1 correspond par exemple à sélectionner une réponse à une question (e.g., sélectionner « Sydney » pour la question « Quelle est la capitale de l'Australie ? ») et la réponse de type-2 correspond au jugement de l'exactitude de cette réponse (e.g., « Je suis certain-e à 50% que j'ai sélectionné la bonne réponse »).

Il existe une grande diversité de jugements métacognitifs pouvant être utilisés et le choix dépend des objectifs de l'étude (voir Figure 1.2). Cette diversité est à prendre en compte lorsque l'on s'intéresse à la métacognition sur la mémoire puisque les questions à se poser quant à leur base, leur précision et les processus les sous-tendant n'auront pas les mêmes réponses selon le type de jugement étudié (e.g., voir Kelemen et al., 2000 pour une étude de corrélations entre plusieurs types de jugements; voir [Chapitre 2](#) pour plus de détails). De plus, leur utilisation dans les activités de la vie quotidienne et leurs conséquences sur la santé mentale dépendent également du type de jugement considéré (Seow et al., 2021). Une première distinction contraste les jugements globaux aux jugements locaux. Les jugements globaux permettent d'étudier la capacité à évaluer la performance générale à une tâche. D'un

point de vue expérimental, les jugements globaux sont généralement une estimation du nombre ou de la proportion d'essais réussis pendant la tâche entière. D'un autre côté, les jugements locaux sont faits sur chaque item testé et permettent une analyse plus fine du monitoring. Les jugements locaux les plus retrouvés dans la littérature sont le jugement de facilité d'apprentissage (*ease-of-learning*, EOL; Underwood, 1966), le jugement d'apprentissage (*judgment-of-learning*, JOL; Ar buckle & Cuddy, 1969), le sentiment de savoir (*feeling-of-knowing*, FOK; Hart, 1965) et les jugements rétrospectifs de confiance (*retrospective confidence judgment*, RCJ; voir Nelson & Narens, 1990).

Une seconde information à prendre en compte lorsque l'on étudie les jugements métacognitifs est le moment auquel ils sont faits. On distingue les jugements prédictifs (i.e., faits avant la réponse de premier ordre, portant sur l'exactitude d'une réponse future) des jugements rétrospectifs (i.e., faits après la réponse, portant sur l'exactitude d'une réponse déjà donnée). Dans une tâche de mémoire, nous pouvons même aller encore plus loin pour départager les jugements prédictifs. Les participants peuvent être amenés à juger leur capacité future à récupérer une information en mémoire au moment de l'encodage des informations (e.g., JOL immédiats), un moment après l'encodage (e.g., JOL différés), ou encore au moment de la tentative de récupération de l'information en mémoire (e.g., FOK).

Alors que les EOL sont des prédictions sur la facilité/difficulté d'apprentissage d'un item et sont émis avant l'encodage de ces items, les JOL et FOK sont faits pendant ou après l'apprentissage du matériel à mémoriser (Nelson & Narens, 1990). Dans un paradigme de EOL, les participants voient les items qu'ils devront apprendre plus tard et doivent évaluer la difficulté qu'ils auront à apprendre ces items (e.g., « à quel point pensez-vous pouvoir apprendre ce mot lors du test ? », Jemstedt et al., 2018).

D'un autre côté, les JOL sont des prédictions sur la capacité future à rappeler ou reconnaître un item. Généralement, les participants apprennent une liste de paires d'items puis les indices (e.g., premier item d'une paire) sont représentés seuls et les participants doivent juger s'ils pensent pouvoir plus tard rappeler/reconnaître chaque cible. Dans les paradigmes de JOL immédiats, les participants font ce jugement immédiatement après avoir encodé chaque paire. Les JOL différés sont quant à eux faits quelques secondes (voire minutes) après l'encodage (Nelson & Dunlosky, 1991). Après un moment, les participants sont ensuite testés sur leur mémoire par un test de rappel libre (e.g., Rhodes & Castel, 2009), indicé (e.g., Dunlosky & Nelson, 1992) ou un test de reconnaissance (e.g., Matvey et al., 2001).

Ensuite, les FOK sont des prédictions sur la reconnaissance future d'un item qui n'est pas récupéré en mémoire (Hart, 1965). Le paradigme le plus classiquement utilisé pour évaluer les FOK est le paradigme de rappel-jugement-reconnaissance. Les participants doivent d'abord rappeler une information spécifique qui a été apprise dans une première phase (i.e., tâche de ME, eFOK) ou qui fait appel à des connaissances générales (i.e., tâche de mémoire

sémantique (MS), sFOK). Lorsqu'ils n'arrivent pas à rappeler l'information cible, ils doivent faire une prédiction sur leur capacité à pouvoir l'identifier parmi des propositions (i.e., jugement de FOK). Enfin, ils doivent reconnaître les cibles parmi des distracteurs. Dans cette thèse, nous avons également utilisé un autre paradigme de prédictions de reconnaissance (*prediction of recognition*, PoR) ressemblant à celui du FOK et des JOL différés dans lequel nous ne demandons pas explicitement aux participants de rappeler l'information cible avant qu'ils ne fassent leur prédiction. Ces PoR sont évaluées sur tous les items et ne suivent pas forcément une tentative de rappel (contrairement aux FOK). Ils ne portent donc pas seulement sur les cibles que les personnes n'arrivent pas à récupérer, et précèdent chaque réponse de type-1.

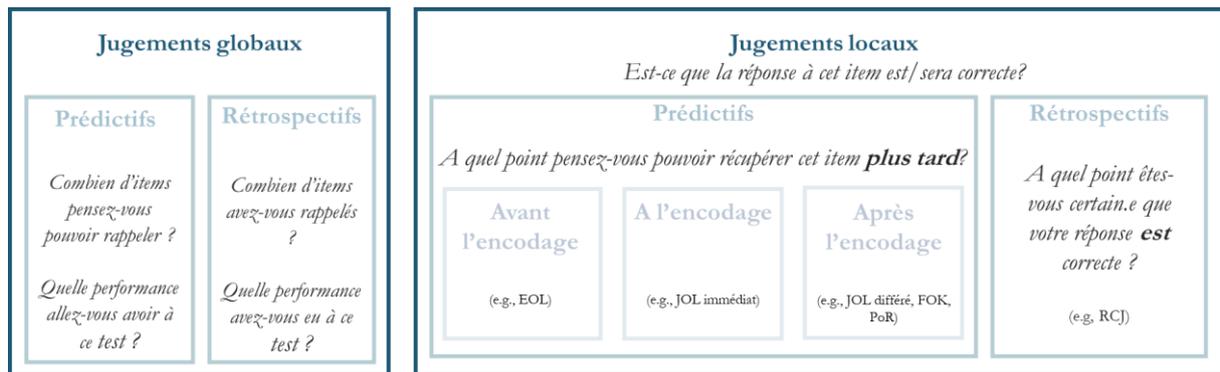
Enfin, les RCJ sont des jugements rétrospectifs émis après que l'information a été récupérée en ME (RCJ épisodique, eRCJ) ou MS (RCJ sémantique, sRCJ), et consiste à juger le degré de confiance en l'exactitude d'une réponse. Ce jugement peut être porté sur n'importe quelle réponse de type-1 comme le rappel libre mais est généralement utilisé dans des tâches de reconnaissance dans lesquelles les participants doivent identifier des informations cibles parmi des distracteurs (e.g., Hart, 1965; Schwartz & Metcalfe, 1992).

Notons que d'autres types de jugements métacognitifs comme les jugements de monitoring de source (Nelson & Narens, 1990), les jugements d'oubli (Finn, 2008; Halamish et al., 2011), les jugements de rétention (Tauber & Rhodes, 2012b), ou encore les jugements d'amélioration (Townsend & Heit, 2011) sont parfois étudiés. Cependant, nous n'en discuterons pas plus dans ce travail car nous nous sommes focalisés sur les jugements de monitoring plus classiques décrits précédemment.

Comme nous le montrerons dans le [Chapitre 2](#), cette distinction entre les différents types de jugements est importante car les questions à se poser quant à leurs bases, aux processus sous-jacents, et leur exactitude dépendent du type de jugement considéré. Dans cette thèse, nous nous focaliserons uniquement sur les jugements locaux et en particulier sur les prédictions de reconnaissance (PoR) dont les FOK font partie (Chapitres 4, 5, 6, 7) et les jugements rétrospectifs (Chapitres 5 et 7).

D'un point de vue comportemental, dans les paradigmes utilisant les mesures de métacognition locales, et notamment rétrospectives, le jugement métacognitif est une décision (de second-ordre) sur une décision (de premier ordre). Ainsi, la décision de premier ordre est la décision mnésique (e.g., choix sur un test de reconnaissance entre 2 alternatives) qui précède ou suit la décision de second ordre qui juge l'exactitude de la décision de premier ordre et constitue le jugement métacognitif. Ces tâches de premier ordre peuvent mettre en jeu la ME (i.e., mémoire des événements vécus dans un contexte spatio-temporel spécifique), la MS (i.e., connaissances générales sur le monde), la mémoire autobiographique (i.e.,

mémoire des évènements et faits relatifs à l'histoire personnelle), ou encore, peut également être une tâche de témoignage oculaire (i.e., identification de l'objet ou de l'auteur d'un crime). Les EOL et les JOL étant faits avant ou pendant l'apprentissage d'un nouveau matériel, ils ne peuvent être étudiés qu'en ME. De plus, les jugements peuvent porter sur la performance à un test de reconnaissance (i.e., choix parmi plusieurs propositions) ou de rappel (i.e., récupération indicée ou libre). Dans cette thèse, nous comparerons les PoR/FOK épisodiques et sémantiques dans les Chapitres 4 et 5. Les Chapitre 6 et 7 n'incluent quant à eux que des PoR/FOK ou RCJ épisodiques.



**Figure 1.2 :** Exemples de jugements métacognitifs étudiés sur des tâches de mémoire (liste non exhaustive).

### 1.2.2. Mesures de monitoring

De nombreuses mesures peuvent être utilisées afin d'évaluer le monitoring (Schraw, 2009, voir Tableau 1.1) et un débat persiste à propos de la meilleure manière de le mesurer (Fleming & Lau, 2014; Goodman & Kruskal, 1954; Masson & Rotello, 2009; Nelson, 1984). D'un côté, la magnitude est la valeur du jugement brut. Cette mesure peut être obtenue en regardant la fréquence des jugements utilisés (e.g., 50% de jugements « oui j'y arriverai ») ou en calculant la moyenne des jugements (e.g., en moyenne la personne donne un jugement de 0.75 sur 1). La magnitude est parfois mentionnée comme étant un biais de réponse de type-2, c'est-à-dire une tendance à utiliser certains points de l'échelle de jugement (e.g., tendance à plus répondre « 80% » ; e.g., Fleming & Lau, 2014; Xue et al., 2021).

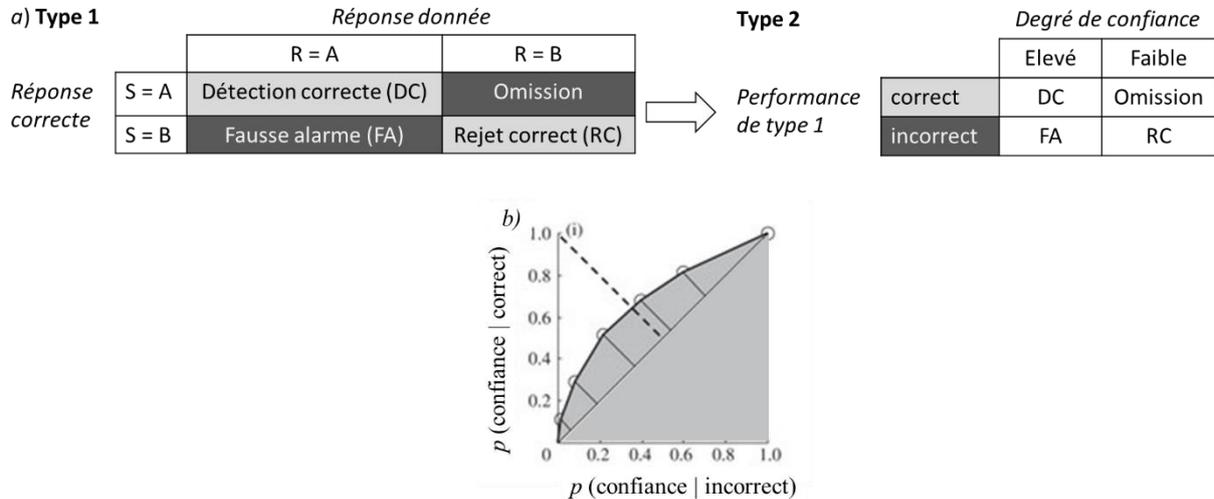
Ensuite, le biais métacognitif renvoie au degré de sur ou sous-estimation de la performance et donne un indice de la direction de l'erreur des jugements. Le biais s'obtient généralement en soustrayant la performance moyenne à la valeur moyenne des jugements métacognitifs. Ainsi, un biais négatif suggère une sous-estimation de la performance alors qu'un biais positif reflète au contraire une surestimation. À noter que pour certains auteurs, le biais métacognitif est plutôt défini comme un biais de réponse de type-2 (e.g., Fleming & Lau,

2014; Xue et al., 2021). Cette conception du biais correspond plus à ce que nous avons appelé magnitude dans ce travail.

D'un autre côté, la sensibilité (ou précision) métacognitive rend compte de la capacité à discriminer les essais corrects des essais incorrects et peut être mesurée dans des paradigmes de jugements locaux uniquement. Dans le domaine de la mémoire, cette précision est généralement mesurée par la corrélation non-paramétrique gamma (Nelson, 1984) qui permet de voir si les jugements métacognitifs reflètent la performance réelle sur la tâche de premier ordre. Une corrélation au-dessus de zéro suggère donc une bonne sensibilité métacognitive. Cependant, cette mesure possède certaines limites. Notamment, le biais de réponse de type-2 (i.e., tendance à donner des jugements élevés ou faibles) est un facteur confondu qui influencerait le score gamma, biaisant l'interprétation de cette mesure (Maniscalco & Lau, 2012; Masson & Rotello, 2009). De plus, cette mesure est aussi influencée par la performance réelle à la tâche de premier ordre puisque des corrélations plus fortes entre les jugements métacognitifs et la performance réelle sont observées lorsque la tâche de premier ordre est mieux réussie (Vuorre & Metcalfe, 2021).

Une autre façon de rendre compte de la sensibilité métacognitive est d'utiliser la logique de la théorie de détection du signal (TDS ou SDT en anglais; Clarke et al., 1959; Green & Swets, 1966, voir Figure 1.3) en calculant par exemple l'aire sous la courbe ROC de type-2 (AUROC2 ; Galvin et al., 2003) ou un  $d'$  de type-2 (Lachman et al., 1979; Nelson, 1984). Ces mesures sont notamment grandement utilisées dans le domaine de la méta-perception, mais moins dans celui de la métamémoire. Cette approche permet d'étudier la capacité à discriminer le signal du bruit (i.e., absence de signal ou signal non pertinent pour la tâche, Macmillan & Creelman, 2004). Sur une tâche de reconnaissance en mémoire, il est ainsi possible de mesurer la performance de type-1 par le score de discrimination  $d'$  en considérant les cibles (i.e., bonnes réponses) comme le signal et les distracteurs (i.e., mauvaises réponses) comme du bruit. La performance de type-2 est également calculée en suivant la même logique (voir Figure 1.3). Il est ensuite possible de tracer une courbe ROC de type-2 cumulative selon la probabilité que le participant ait donné un certain niveau de confiance (ou moins) sachant que la réponse était correcte (en ordonnée) et la probabilité qu'il ait donné au plus ce même niveau de confiance sachant que la réponse était incorrecte (en abscisse). Plus l'AUROC2 est élevée, plus la sensibilité métacognitive est forte. Une AUROC2 de 0.5 indique une faible sensibilité et des réponses de type-2 équivalentes au hasard (diagonale sur la Figure 1.3b). L'avantage de ces mesures basées sur la TDS est la moindre influence du biais de réponse de type-2 (mais voir Azzopardi & Evans, 2007). L'utilisation du  $d'$  de type 2 nécessite plusieurs prérequis (e.g., normalité des distributions du signal et du bruit, et variance égale). Cependant, ces prérequis ne sont pas toujours respectés et ils ne sont pas nécessaires pour l'utilisation des corrélations non-paramétriques gamma et des AUROC2 (Masson &

Rotello, 2009). De plus, l'AUROC2 et le  $d'$  de type-2 sont également influencés par la performance à la tâche de premier ordre (Vuorre & Metcalfe, 2021).



**Figure 1.3** : Logique de la TDS.

- a) Tableaux de contingence pour la TDS de type-1 et de type-2.  
 b) Courbe ROC de type-2. L'aire sous la courbe (AUROC2) est représentée par la partie grise.  
 Traduits et adaptés de Fleming & Dolan (2012) et Fleming & Lau (2014).

Pour répondre aux limites des mesures de sensibilité métacognitive citées précédemment (i.e., influence du biais de réponse et de la performance de premier ordre), des mesures d'efficacité métacognitive ont été développées (Maniscalco & Lau, 2012). Il s'agit d'une manière de calculer la sensibilité métacognitive en contrôlant la performance de premier ordre. Se basant aussi sur la théorie de détection du signal, les auteurs ont proposé un modèle permettant d'évaluer cette sensibilité de manière indépendante de la magnitude (meta- $d'$ ; Fleming & Lau, 2014; Maniscalco & Lau, 2012). Ce meta- $d'$  correspond à la performance de type-1 qu'un observateur « idéal » (i.e., qui serait métacognitivement parfait) aurait et il est estimé sur la base des données de type-2. L'efficacité métacognitive est ensuite évaluée en estimant le ratio, nommé M-ratio, entre le meta- $d'$  et le  $d'$  de type-1 (i.e., la performance réelle).

En résumé, les mesures de sensibilité et d'efficacité métacognitives possèdent chacune des avantages et des inconvénients. Dans ce travail de thèse, nous avons utilisé la méthode corrélationnelle très largement répandue dans le domaine de la métamémoire ainsi que les mesures moins biaisées se basant sur la TDS telles que le  $d'$  de type-2 et les courbes ROC de type-2 (voir . Nous avons également essayé d'estimer l'efficacité métacognitive dans les Chapitres 6 et 7, mais ces analyses ne se sont révélées concluantes (car sûrement plus pertinentes) que dans le [Chapitre 7](#). Le but n'était pas de comparer ces mesures mais plus de présenter des résultats complémentaires.

**Tableau 1.1** : Mesures de monitoring décrites dans la littérature.

Facette du monitoring	Explication	Autres termes utilisés	Exemples de mesures
<b>Magnitude</b>	Valeur du jugement sur l'échelle (e.g., 50%)	Jugement moyen Biais de réponse de type-2	Moyenne ou proportion des jugements
<b>Biais</b>	Degré de sur- ou sous-estimation	Calibration Précision absolue	Différence entre performance réelle et magnitude des jugements (surestimation : biais > 0 ; sous-estimation : biais < 0) M-distance
<b>Sensibilité</b>	Capacité à discriminer les essais corrects des essais incorrects	Précision relative Résolution Discrimination	Corrélation gamma d' de type 2 Aire sous la courbe ROC de type-2 Meta-d'
<b>Efficiace</b>	Sensibilité contrôlée pour le biais et la performance mnésique		Meta-d'/d' = M ratio

L'ensemble des mesures présentées ici peut être étudié sur la majorité des tâches utilisant les jugements métacognitifs locaux précédemment décrits. Elles ne sont pas forcément affectées par les mêmes facteurs et sont donc complémentaires (Higham et al., 2016 ; voir par exemple Murphy et al., 2022 pour l'exemple d'un effet d'une manipulation sur la magnitude des jugements sans effet significatif sur le biais ou la sensibilité métacognitive). Si comme expliqué précédemment, ces mesures sont reliées entre elles au niveau statistique, elles sont indépendantes d'un point de vue théorique. En effet, une personne ayant tendance à se sous-estimer pourrait quand même être sensible aux variations de sa performance d'un essai à l'autre. Elles permettent ainsi de mesurer différents aspects de la métacognition (e.g., degré de sur/sous-estimation ou sensibilité aux changements de performance). L'utilisation de ces mesures dépend des intérêts de l'étude. La magnitude, le biais, la sensibilité et l'efficiace seront utilisés comme mesures de métacognition dans ce projet de thèse. D'autres mesures de la métacognition telles que la discrimination et la dispersion ont également été décrites (Schraw, 2009).

### 1.3. Conclusion du Chapitre 1

La métacognition est une capacité de haut-niveau hétérogène. Cette hétérogénéité se traduit par la diversité des composants métacognitifs (connaissances, régulation), des processus (monitoring ou contrôle), des jugements de monitoring (JOL, FOK, RCJ) mais aussi des mesures (magnitude, biais métacognitif, sensibilité).

L'efficacité de la mémoire repose sur des processus métacognitifs intacts qui guident les apprentissages et la récupération des informations en mémoire bien qu'elle ne soit pas nécessaire à toute capacité mnésique (Dunlosky & Bjork, 2014). Plusieurs questions se posent vis-à-vis des capacités métacognitives. En particulier, nous essayerons de mieux comprendre sur quoi se basent les jugements métacognitifs, les processus les sous-tendant et les raisons pour lesquelles ils ne sont parfois pas aussi précis et corrects.

## CHAPITRE 1 : POINTS IMPORTANTS

- Ce travail porte sur la métacognition sur des tâches de **mémoire** (i.e., métamémoire).
  - La métacognition se manifeste de plusieurs manières : connaissances, régulation.
- Plus spécifiquement, l'intérêt de cette thèse sera porté sur le processus de **monitoring** et non celui de contrôle.
- Plusieurs jugements de monitoring peuvent être étudiés. Notamment, les études de cette thèse comportent des jugements **locaux** ou **globaux**, **prédictifs** ou **rétrospectifs**, **épisodiques** ou **sémantiques**.

# Chapitre 2.

## Bases des jugements métacognitifs

---

---

Les travaux dans le domaine de la métacognition ont évalué la capacité des individus à dissocier leur aptitude à rappeler une information plutôt qu'une autre, et à juger que l'un de leurs souvenirs est correct mais pas un autre. Ces travaux ont donc cherché à identifier les bases de ces jugements métacognitifs. De plus, d'autres études ont tenté d'évaluer si un jugement sur une tâche est équivalent à un autre. Plusieurs théories complémentaires ont été proposées et seront discutées dans ce chapitre.

### 2.1. Sources de l'auto-évaluation

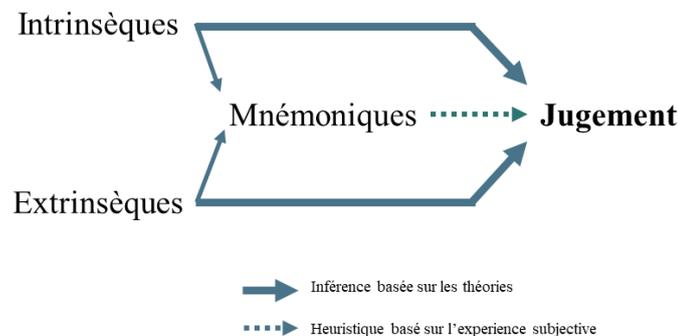
#### 2.1.1. Théorie de l'utilisation d'indices : présentation générale

L'hypothèse d'accès direct (*direct-access hypothesis* en anglais, Hart, 1967; voir commentaires dans Nelson et al., 1984; Schwartz, 1994) a été proposée pour expliquer l'exactitude des jugements prédictifs et en particulier le FOK. Selon cette hypothèse, la capacité future à récupérer une information en mémoire est évaluée sur la base d'une estimation directe de la force de la trace mnésique. Cette théorie suppose donc que les personnes ont un moniteur interne permettant d'évaluer directement la présence, l'accessibilité et la force de cette trace mnésique à tout moment. Plus la force d'un souvenir est importante, plus les participants vont prédire des performances élevées sur l'essai correspondant. Cependant, selon cette hypothèse, les personnes devraient donc être sensibles aux facteurs influençant la force d'un souvenir, et toute influence sur la force de la trace devrait avoir un effet sur la magnitude du jugement, ce qui n'est pas toujours le cas (e.g., Carroll et al., 1997 ; voir Schwartz, 1994). Par exemple, Carroll et al. (1997) ont réussi à faire en sorte que certaines paires d'items soient mieux retenues que d'autres par les participants, tout en montrant que cette manipulation n'entraînait pas nécessairement les participants à prédire de meilleures performances pour les paires mieux retenues (cette étude sera détaillée plus en détails p. 50). De plus, cette théorie ne permet pas d'expliquer les erreurs de jugements (e.g., estimer qu'une information peut être récupérée alors que ce n'est pas le cas). Ainsi, cette hypothèse a depuis été remise en question (Reder & Ritter, 1992; Schwartz & Metcalfe, 1992) et des théories alternatives ont été proposées, comme la théorie d'utilisation d'indices (*cue-utilization view* en anglais; Koriat, 1997). Ces théories fournissent un cadre commun aux jugements prédictifs et rétrospectifs.

Selon la théorie de l'utilisation d'indices (Koriat, 1997), les jugements métacognitifs sont par nature inférentiels, c'est-à-dire que les personnes utilisent des indices auxquels elles sont sensibles pour inférer une estimation sur l'exactitude de leur réponse. Initialement développée pour les JOL, cette théorie a ensuite été étendue aux autres jugements tels que les FOK et les RCJ. Plusieurs sources de ces indices ont été identifiées. Premièrement, ces indices peuvent être basés sur la théorie (*theory-based*; i.e., connaissances/croyances métacognitives ; e.g., penser qu'un mot écrit en plus gros est plus facile à apprendre qu'un mot écrit en plus petit) ou sur l'expérience (*experience-based*; e.g., avoir l'impression subjective qu'un mot est mieux mémorisé qu'un autre) (Koriat et al., 2008; Koriat & Levy-Sadot, 1999). Les croyances, ou théories « naïves », influenceraient les jugements de manière analytique via une inférence alors que les sensations ou expériences métacognitives ne seraient pas analytiques mais auraient plutôt une influence plus implicite sur les jugements métacognitifs (Koriat, 1997). Cette distinction a été inspirée par les théories des doubles processus qui proposent deux modes de pensée : un mode rapide et automatique (i.e., heuristique basé sur les expériences subjectives) et un mode plus lent et délibéré (i.e., inférences basée sur la théorie) (Koriat et al., 2008).

De plus, différentes informations peuvent être transmises par ces indices : extrinsèques, intrinsèques, ou mnémoniques (Koriat, 1997). Les indices extrinsèques sont relatifs aux conditions de passation ou aux stratégies d'encodage mises en place par le participant. Par exemple, dans une tâche de mémoire, un participant peut utiliser le nombre de fois qu'un item a été présenté dans la phase d'apprentissage ou le temps de présentation comme indice pour évaluer sa performance. Ensuite, les indices intrinsèques correspondent aux caractéristiques des items à apprendre. Il s'agit des a priori ou sensations vis-à-vis de la facilité d'apprentissage des stimuli. Des exemples de ce type d'indices sont la force d'association perçue entre deux items lors de l'apprentissage de paires de mots (e.g., voir Rabinowitz et al., 1982 pour un exemple dans le JOL), et le degré d'imagerie ou la fréquence d'un mot (e.g., Bruce et al., 1982). Enfin, les indices mnémoniques sont des sensations subjectives internes au sujet vis-à-vis du traitement de l'information à l'encodage ou au test (e.g., facilité de traitement ; Benjamin et al., 1998 ; accessibilité à des informations pertinentes, Koriat, 1993 ; voir Koriat, 1997 pour d'autres exemples). Ces indices mnémoniques peuvent être influencés par les indices extrinsèques et intrinsèques. Ainsi, les indices extrinsèques et intrinsèques peuvent directement influencer les jugements métacognitifs via l'application d'une théorie naïve, une croyance, ou indirectement via les indices mnémoniques et donc des indices basés plus sur l'expérience (voir Figure 2.1). Un exemple d'influence indirecte est la grande taille d'un stimulus (i.e., indice intrinsèque) qui va produire une sensation de facilité d'encodage (i.e., indice mnémonique) amenant un participant à penser pouvoir plus rappeler cet item qu'un stimulus plus petit (Undorf et al., 2017).

Ces indices sont corrects, ou diagnostiques, lorsqu'ils correspondent à des facteurs qui influencent réellement la performance de premier ordre. Cependant, dans certains cas, ils ne sont pas diagnostiques, c'est-à-dire qu'ils ne sont pas prédictifs de la performance objective à la tâche de premier ordre.



**Figure 2.1** : Types d'indices à la base des jugements métacognitifs selon Koriat (1997).

### 2.1.2. Fluence de traitement : influence sur les JOL, les FOK, et les RCJ

De nombreux indices ont été documentés dans la littérature de la métamémoire. Les différents types de fluence font partie des indices les plus étudiés. La fluence peut être définie par la facilité subjective avec laquelle une information est traitée (Jacoby & Dallas, 1981; Oppenheimer, 2008). Plusieurs types de fluence ont été décrits en lien avec les jugements métacognitifs comme la fluence perceptive, la fluence de récupération, et la fluence conceptuelle (Alter & Oppenheimer, 2009). La fluence perceptive est liée aux caractéristiques physiques des stimuli. Par exemple, la magnitude des JOL est particulièrement influencée par les modifications du volume des sons (e.g., prédictions de rappel plus élevées pour les sons plus forts, Expérience 2 de Murphy et al., 2022; Rhodes & Castel, 2009), le degré de floutage d'un mot (e.g., prédictions de rappel et de reconnaissance plus faibles pour les mots flous, Yue et al., 2013), la taille de police d'écriture (e.g., prédictions plus élevées pour des écritures larges; McDonough & Gallo, 2012; Rhodes & Castel, 2008), le masquage de mots (Besken & Mulligan, 2013), ou encore la fragmentation d'images (Besken, 2016).

La fluence de récupération est la facilité avec laquelle une information est récupérée en mémoire et est généralement mesurée par les temps de réponse sur une tâche de rappel ou de reconnaissance (e.g., Kelley & Lindsay, 1993; Koriat, 1993; Zakay & Tuvia, 1998). Plusieurs études mettent en avant une influence de la fluence de récupération sur les RCJ : les personnes perçoivent les réponses données plus rapidement comme plus correctes que des réponses données plus lentement, qu'elles le soient réellement ou pas. Ce résultat a été observé sur une tâche de MS testant les connaissances générales (e.g., Kelley & Lindsay, 1993) et sur une tâche de décision perceptive (e.g., Grimaldi et al., 2015). De plus, d'autres

études ont mis en avant l'influence de la fluence de réponse au moment de l'encodage d'une information sur les prédictions de rappel libre (Expérience 1 de Benjamin et al., 1998) ou de reconnaissance (Matvey et al., 2001). Par exemple, dans une étude (Expérience 1 de Benjamin et al., 1998), le temps que les participants mettaient pour répondre à des questions de culture générale était mesuré. Ils devaient ensuite prédire la probabilité qu'ils puissent par la suite se souvenir des réponses. Après 20 minutes, ils devaient rappeler les réponses qu'ils avaient données, sans que les questions soient représentées. Les participants prédisaient rappeler plus de réponses trouvées rapidement dans le test initial (i.e., sorte de phase d'encodage) que des réponses plus lentes, malgré qu'ils aient été meilleurs pour rappeler les réponses qui leur étaient venues plus lentement. Dans une autre étude (Hertzog et al., 2003), la fluence de récupération était également mesurée au moment de l'encodage (les auteurs utilisaient d'ailleurs le concept de fluence d'encodage) mais correspondait à la facilité avec laquelle les participants arrivaient à former une image mentale interactive entre les deux mots d'une paire et non la facilité à récupérer l'information cible. La magnitude des JOL était plus élevée dans les essais dans lesquels les participants avaient réussi à former rapidement une image au moment de l'encodage. Ainsi, les exemples précédents montrent une influence de la fluence de récupération sur la magnitude des jugements. La fluence de réponse à l'encodage peut également jouer sur la sensibilité métacognitive des JOL (Matvey et al., 2001).

Enfin, la fluence peut être manipulée à travers des facteurs sémantiques : on parle alors de fluence conceptuelle. Ce type de fluence sera l'objet du [Chapitre 7](#). Ce type de fluence dépend des connaissances générales que les personnes ont sur le monde (i.e., MS) et est généralement plus élevé dans des cas où l'information est traitée dans un contexte congruent avec ces connaissances. Une manière de manipuler la fluence conceptuelle est par le degré d'association sémantique entre deux mots d'une paire à apprendre ou entre les mots d'une liste. Cette association sémantique influence notamment les JOL (e.g., Carroll et al., 1997; Dunlosky & Matvey, 2001; Matvey et al., 2006; Undorf & Erdfelder, 2015; voir Mueller et al., 2013 pour une revue pour les JOL). Plus particulièrement, les participants prédisaient une meilleure mémorisation des paires reliées (e.g., chaise – table) par rapport aux paires non-reliées (e.g., chien – cuillère), même lorsque les performances étaient en réalité meilleures pour les paires non-reliées (Carroll et al., 1997). Dans cette même étude, le degré d'association était moins pris en compte que le degré d'apprentissage lorsque le jugement étudié était un FOK et donc que la prédiction était faite au moment de la récupération (Carroll et al., 1997). L'effet de la fluence conceptuelle sur les jugements métacognitifs peut également être manipulée par l'effet de fréquence des mots sur les JOL (Benjamin, 2003). Dans cette étude, les participants prédisaient de meilleures capacités à reconnaître les mots fréquents par rapport aux mots plus rares, bien qu'ils aient en réalité reconnu plus de mots rares. D'autres exemples de fluence conceptuelle seront également discutés dans le [Chapitre 3](#).

Ainsi, lorsque les manipulations de la fluence de traitement influencent la magnitude des jugements métacognitifs, elles n'entraînent pas systématiquement une modification de la performance en mémoire (e.g., Expérience 2 de Besken & Mulligan, 2013; Rhodes & Castel, 2009; Expériences 2b et 3 de Yue et al., 2013) ou peuvent entraîner un effet inverse à celui prédit par les participants (e.g., Expérience 1 de Besken & Mulligan, 2013; Carroll et al., 1997). Ces modifications peuvent donc créer des illusions métacognitives. De plus, bien qu'étant une sensation subjective, certains auteurs proposent que la fluence n'influence pas les jugements métacognitifs directement mais plutôt à travers un processus d'attribution et le développement de théories naïves (Hu et al., 2015; Mueller et al., 2014 ; mais voir Yang et al., 2018 pour des résultats contradictoires). En effet, les personnes infèreraient des causes et des conséquences à cette fluence (Oppenheimer, 2008). L'exemple le plus connu est l'effet de police d'écriture sur les JOL (Murphy et al., 2022; Rhodes & Castel, 2008; Yang et al., 2018). Dans ces études, les participants apprenaient des mots écrits en petite ou large police puis prédisaient leur capacité à rappeler ces mots. Globalement, les participants prédisaient moins rappeler de mots présentés en petite plutôt qu'en large écriture bien que le rappel ne soit pas réellement affecté par cette manipulation. La manipulation de la taille d'écriture pourrait jouer sur la fluence perceptive, et cette fluence influencerait sur les jugements métacognitifs via des processus basés sur l'expérience (Besken & Mulligan, 2013; Undorf et al., 2017). Cependant, selon d'autres auteurs, cet effet serait en partie dû à la croyance que cette manipulation a une influence sur la mémoire (Mueller et al., 2014). D'autres travaux mettent en évidence que les deux types de processus joueraient un rôle : la fluence serait utilisée comme indice à la fois à travers les processus basés sur les théories et ceux basés sur les expériences (Besken, 2016; Undorf & Erdfelder, 2015).

### **2.1.3. Familiarité avec l'indice et récupération partielle de la cible : sources du FOK et des JOL différés**

Les études portées sur le FOK (i.e., prédiction d'une reconnaissance future d'information temporairement non récupérable) ont prouvé que même lorsque l'information cible n'est pas récupérée, les personnes peuvent quand même évaluer si elle pourra être accessible plus tard. Ces études ont montré que deux principaux facteurs influencent ces prédictions : la familiarité subjective avec l'indice (i.e., l'item d'une paire dans une tâche épisodique, ou la question de culture générale dans une tâche sémantique; Metcalfe et al., 1993; Reder, 1988; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992) et les informations partielles accessibles sur la cible à récupérer (Koriat, 1993, 1994, 1995; Koriat et al., 2003). Initialement mis en concurrence, un modèle interactif propose que ces deux théories sont complémentaires (Koriat & Levy-Sadot, 2001). En effet, les FOK seraient initialement influencés par un

processus rapide de familiarité avec la question posée ou l'indice affiché. De cette manière, les participants formeraient une première estimation de la probabilité qu'ils reconnaissent la cible. Ensuite, s'ils sont assez familiers avec l'indice, les personnes peuvent s'engager dans une recherche active des informations à propos de la cible à récupérer. Ainsi, dans les cas où la cible n'est pas récupérée en entier, des informations partielles peuvent tout de même être accessibles et guider les FOK. Ces informations peuvent notamment être de nature orthographique (e.g., « ça commence par un g »), phonologique (e.g., « ça rime avec télégraphe ») ou sémantique (e.g., « c'est un animal »).

L'influence de la familiarité avec l'indice dans les paradigmes de FOK a été mise en évidence par plusieurs travaux. Par exemple, la magnitude des FOK étaient plus élevés lorsque les indices avaient été présentés avant la phase de prédiction (i.e., indices amorcés) que lorsque les indices n'étaient pas amorcés. Cet effet était retrouvé que l'indice soit les termes d'une question de culture générale (i.e., tâche de FOK sémantique, sFOK; Reder, 1987) ou l'indice d'une paire de mots (i.e., tâche de FOK épisodique, eFOK ; Schwartz & Metcalfe, 1992). Ainsi, rendre plus familier un indice en le présentant avant que les participants ne doivent faire le FOK augmentaient leur sensation de savoir et de pouvoir par la suite retrouver la réponse. En ce qui concerne le rôle de l'accessibilité à la cible, d'autres études (Koriat, 1993; Thomas et al., 2012) ont montré que la quantité d'informations récupérées au moment de la tentative de rappel était importante puisque la magnitude et la sensibilité des FOK augmentaient avec le nombre d'informations partielles récupérées, indépendamment de l'exactitude réelle de ces informations (sauf dans le cas où ces informations étaient sémantiques). En effet, la qualité des attributs sémantiques récupérés mais non celle des attributs perceptifs influençait également la magnitude des FOK puisque les FOK étaient plus élevés dans le cas où les informations sémantiques partielles étaient correctes (Thomas et al., 2012).

De même, l'effet de la familiarité avec l'indice présenté au moment du jugement et l'accessibilité de la cible a également été retrouvé sur les JOL différés (Metcalfe & Finn, 2008a). Ainsi, ces deux types d'indices semblent particulièrement affecter les prédictions faites un moment après l'encodage.

#### **2.1.4. Remarques additionnelles et conclusions sur la théorie de l'utilisation d'indices**

Si nous avons montré que les jugements métacognitifs se basent sur différents indices, l'influence de ces indices dépend des conditions de passation. Notamment, certains indices influencent les jugements lorsqu'ils sont manipulés dans un design intra-sujets mais pas dans un design inter-sujets (e.g., Carroll & Nelson, 1993; Yue et al., 2013). Par exemple, l'effet de

sur-apprentissage observé sur les eFOK n'était significatif que dans les cas où les mêmes participants étudiaient des items selon chacun des deux critères (Carroll & Nelson, 1993). Dans cette étude, les participants devaient apprendre des paires d'items puis ils devaient rappeler les cibles sur présentation des indices. Plusieurs critères d'apprentissage étaient fixés : les participants devaient réapprendre les paires jusqu'à ce qu'ils puissent rappeler une fois la cible, mais pour d'autres paires (design intra) ou d'autres participants (design inter), ce critère était fixé à 6 rappels non-consécutifs. Quatre semaines plus tard, les participants étaient testés sur leur mémoire et ils devaient faire un FOK pour les cibles qu'ils n'arrivaient pas à rappeler. Les participants prédisaient une meilleure reconnaissance des items appris avec un critère de 6 rappels (i.e., paires sur-apprises) par rapport aux paires rappelées une seule fois (i.e., paires sous-apprises), mais uniquement lorsque le critère était manipulé en intra-sujets. Cet effet de sur-apprentissage n'était pas reproduit lorsque le critère de rappel était manipulé à travers plusieurs groupes. Ces études suggèrent également que les jugements sont comparatifs : la magnitude des jugements métacognitifs change selon les différentes modalités des facteurs manipulés mais principalement lorsque le participant fait l'expérience des différentes conditions.

De plus, plusieurs travaux ont montré un lien entre plusieurs indices comme la familiarité et la fluence (Whittlesea, 1993; Whittlesea et al., 1990). Dans des contextes écologiques, il est donc probable que plusieurs indices jouent sur une évaluation métacognitive, et qu'à la fois des indices basés sur la théorie et d'autres basés sur l'expérience contribuent à la formation de ces jugements métacognitifs. Des travaux ont, par exemple, mis en évidence que les individus intégraient parfois plusieurs indices lorsqu'ils faisaient un jugement métacognitif (Undorf et al., 2018; Undorf & Bröder, 2020).

Ainsi, bien que tous les types de jugements métacognitifs semblent être inférés sur la base d'indices, les travaux présentés précédemment suggèrent que chaque indice n'influence pas tous les types de jugements de la même manière, et que les personnes n'utilisent pas les mêmes types d'indices selon le moment où le jugement est fait. Par exemple, la fluence de récupération a été identifiée comme servant d'indice davantage pour les jugements rétrospectifs, et la fluence à l'encodage a été particulièrement étudiée dans les paradigmes de JOL (Undorf & Erdfelder, 2015). Selon le moment où le jugement est fait, différentes informations sont utilisées. Est-ce que cela signifie que tous les jugements métacognitifs ne sont pas égaux ? De plus, une grande majorité des études sur la fluence ont porté sur les JOL mais encore peu d'études se sont focalisées sur la fluence (notamment conceptuelle) dans les prédictions de reconnaissance faite au moment de la récupération en mémoire et dans les jugements rétrospectifs. Enfin, quelques études ont montré l'utilisation d'attributs sémantiques dans le FOK, mais des explorations complémentaires sont nécessaires.

## 2.2. Généralité versus spécificité de la métacognition

La question de la généralité des capacités métacognitives est un débat d'actualité. Plusieurs travaux ont eu pour objectif de déterminer si la capacité d'un individu à évaluer sa performance sur une tâche s'étendait à une autre tâche. Par exemple, une question importante est de savoir si une personne qui prédit de manière exacte ses performances arrive aussi bien à les évaluer rétrospectivement. Plusieurs études ont ainsi cherché à savoir si les mêmes processus étaient mis en jeu dans tous les jugements métacognitifs. Cette question se pose vis-à-vis des domaines étudiés dans la tâche de premier ordre (e.g., tâche perceptive ou tâche de mémoire ; voir Ais et al., 2016; Fitzgerald et al., 2017; Mazancieux et al., 2020; McCurdy et al., 2013; Morales et al., 2018 pour des exemples), mais également à l'intérieur d'un domaine cognitif (e.g., dans la méta-perception; Faivre et al., 2018; Song et al., 2011). Bien que plusieurs cadres théoriques communs à tous les jugements ait été proposé pour expliquer qu'il puisse y avoir des mécanismes sous-jacents communs à plusieurs jugements (e.g., *isomechanism framework*, Dunlosky & Tauber, 2014 ; *Cognitive Awareness Model*, Agnew & Morris, 1998; Hannesdottir & Morris, 2007; Morris & Mograbi, 2013), des preuves comportementales, neuropsychologiques et neuro-fonctionnelles suggèrent des processus divergents entre les jugements, notamment en ce qui concerne la sensibilité métacognitive. En effet, plusieurs études laissent penser que le biais métacognitif est domaine-général (Ais et al., 2016; Baranski & Petrusic, 1995; Fleming et al., 2016; Mazancieux, Dinze, et al., 2020). Cela suggère qu'une personne qui a tendance à se surestimer sur une tâche, se surestime également sur une autre tâche. Cependant, plusieurs résultats dont nous allons maintenant discuter montrent que la sensibilité métacognitive semble plus dépendante de la tâche de premier ordre.

### 2.2.1. Preuves comportementales

Certaines études comparant directement les performances métacognitives sur des tâches mettant en jeu différents types de jugements suggèrent une spécificité des types de jugements. Des auteurs ont analysé les corrélations entre les performances métacognitives sur des tâches utilisant plusieurs types de jugements. Par exemple, dans une étude (Leonesio & Nelson, 1990), les participants réalisaient une tâche de mémorisation de mots et faisaient des EOL, un jugement d'apprentissage (estimer à quel point ils connaissaient la réponse) et un FOK. La sensibilité métacognitive, mesurée par des scores gammas, était significativement corrélée entre ces trois types de jugements mais faible (entre 0.12 et 0.19). Les auteurs ont ainsi proposé que ces jugements reposaient sur des « structures sous-jacentes multidimensionnelles ». À noter que ces jugements étaient faits pour les mêmes items et sur la même tâche donc ils n'étaient pas indépendants l'un de l'autre et ont pu s'influencer. Dans

une autre étude (Kelemen et al., 2000), la sensibilité des JOL immédiats et des FOK n'étaient pas significativement corrélée. Cependant, dans cette étude, les JOL étaient mesurés sur une tâche de ME alors que les FOK étaient faits sur une tâche de MS. Comme nous en discuterons dans la suite de ce chapitre, cette différence de mémoire testée sur la tâche de premier ordre peut être un facteur confondu à ce manque de corrélation. Enfin, dans une étude réalisée chez des personnes âgées et plus jeunes (Souchay & Isingrini, 2012), la sensibilité métacognitive des deux groupes était meilleure pour les JOL que pour les FOK, et encore plus chez les personnes âgées. De plus, les deux types de jugements n'étaient pas corrélés. En résumé, ces trois études suggèrent que les capacités métacognitives diffèrent, même d'un jugement prédictif à un autre.

D'autres travaux suggèrent que différents mécanismes sont en œuvre selon la temporalité des jugements. Dans une autre étude (McDonough et al., 2021), les participants réalisaient plusieurs blocs d'apprentissage de paires d'items. Selon le bloc, ils devaient faire des EOL, des JOL, des FOK ou des RCJ. Les analyses ont révélé qu'il ne semblait pas y avoir une compétence unique commune à tous les jugements de monitoring. Plutôt, les auteurs proposent qu'il y aurait deux compétences : une sous-tendant les jugements faits au moment de l'encodage (EOL, JOL), et une sous-tendant les jugements faits au moment de la récupération (FOK, RCJ). Des divergences ont également été observées en comparant des jugements prédictifs aux jugements rétrospectifs. Par exemple, les RCJ faits pour une tâche de rappel étaient plus prédictifs du rappel futur que les JOL (Dougherty et al., 2005; voir Dougherty et al., 2018 et Putnam et al., 2022 pour des répliques). Dans ces études, les participants apprenaient d'abord des paires de mots puis, après un moment, ils réalisaient une tâche de rappel indicé. Après chaque tentative de rappel, ils devaient estimer à quel point ils étaient certains d'avoir rappelé la bonne paire (RCJ) et/ou s'ils pensaient pouvoir retrouver la cible 10 minutes plus tard (JOL). Enfin, ils refaisaient un test final de rappel indicé après un délai de 10 minutes. La sensibilité métacognitive, évaluée en comparant les RCJ et les JOL à la performance de rappel final par une corrélation gamma, était plus élevée pour les RCJ que pour les JOL. De plus, une plus grande sensibilité des jugements rétrospectifs par rapport aux jugements prédictifs a été observée dans des tâches ne mesurant pas la mémoire. Ces résultats ont entre autres été trouvés dans une tâche de détection d'anagrammes (Siedlecka et al., 2016) et dans une tâche de discrimination visuelle (Fleming et al., 2016). Dans ces études, les jugements faits a posteriori de la tâche de premier ordre étaient plus précis que les jugements faits a priori. Ces résultats suggèrent que les indices sur lesquels les participants se basent pour évaluer rétrospectivement leur performance sont plus diagnostiques de la performance réelle que les indices utilisés pour prédire leur performance.

### 2.2.2. Preuves neuropsychologiques et neuro-fonctionnelles

En complément des résultats précédents, des dissociations de sensibilité métacognitive selon le type de jugements ont été observées dans certaines populations. Notamment, sur des tâches épisodiques, les personnes âgées semblent aussi sensibles que des personnes plus jeunes pour prédire leurs performances en rappel dans un paradigme de JOL (e.g., Connor et al., 1997; Dunlosky & Hertzog, 2000; Emanuel Robinson et al., 2006), ou pour évaluer rétrospectivement leurs performances dans un paradigme de RCJ (e.g., Dunlosky & Hertzog, 2000). De plus, une autre étude n'a pas montré de corrélation significative entre l'âge et l'efficacité métacognitive de RCJ sur une tâche de ME (Palmer et al., 2014). Cependant, les personnes âgées présentent des difficultés à prédire leur capacité à reconnaître un item non rappelé sur une tâche de eFOK (Sacher et al., 2015; Souchay et al., 2000), bien que cet effet de l'âge sur le FOK ne soit pas toujours retrouvé et soit débattu (Eakin et al., 2014; Eakin & Hertzog, 2012; Hertzog et al., 2010, voir [Chapitre 4](#) pour une revue). Cette dissociation a été observée dans une étude dans laquelle les mêmes personnes âgées et jeunes ont réalisé une tâche de JOL différé et une de eFOK (Souchay & Isingrini, 2012). Un déficit de sensibilité métacognitive lié à l'âge était observé sur la tâche de eFOK (bien que cet effet disparût lorsque les performances de rappel étaient prises en compte dans les analyses) mais pas sur celle de JOL.

Un même pattern de résultats a été observé chez des patients atteints de MA. En effet, ces patients présentaient un déficit de sensibilité métacognitive sur des tâches de eFOK (Souchay et al., 2002) mais pas sur des tâches de JOL ou eRCJ (Moulin et al., 2000, 2003; Pappas et al., 1992). Ce même pattern de performances a également été observé chez des personnes présentant un trouble du spectre autistique (Grainger et al., 2014, 2016; Wojcik et al., 2013, 2014).

Enfin, des patients présentant des lésions des lobes frontaux ne montraient pas de difficulté à évaluer rétrospectivement leur capacité de rappel sur une tâche de ME dans laquelle ils devaient mémoriser le dernier mot de phrases, bien qu'ils aient montré des eFOK (i.e., prédictions) moins précis par rapport à un groupe contrôle sur cette même tâche (Schnyer et al., 2004). Par contre, d'autres patients avec atrophies frontales latérales présentaient un pattern de performances inverse sur des tâches de MS de reconnaissance de personnes célèbres puisque leur sensibilité était altérée pour des sRCJ mais pas pour des sFOK (Pannu et al., 2005).

La question de la spécificité des jugements métacognitifs peut également trouver des réponses grâce aux études de neuroimagerie fonctionnelle (IRMf). Plusieurs études mettent en avant des réseaux cérébraux communs aux différents jugements métacognitifs mais

également des réseaux plus spécifiques (pour des revues, voir Chua et al., 2014; Fleming & Dolan, 2012). En particulier, la sensibilité métacognitive des JOL, FOK et RCJ est reliée à l'activation des régions frontales antérieures, dans les parties plus médiales pour les prédictions (JOL et FOK) et plus latérales pour les RCJ. De plus, une autre étude a mis en avant des activations différentes en comparant des FOK à des RCJ sur une tâche de ME (Chua et al., 2009). Dans cette étude, les participants apprenaient des associations visages-noms. Ils étaient ensuite testés sur leur mémoire des noms lorsque chaque visage était représenté et faisaient d'abord un FOK, puis une reconnaissance en choix forcé et enfin, un jugement de confiance (RCJ). Les RCJ étaient plus reliés à une activation dans le cortex préfrontal antérieur gauche que les FOK. À l'inverse, les FOK activaient entre autres plus l'hippocampe, le cortex temporal supérieur droit, et le gyrus fusiforme que les RCJ. Les auteurs proposent que ceci pourrait s'expliquer par une plus grande implication des processus de récupération en mémoire dans les FOK par rapport aux RCJ.

Une méta-analyse récente (Vaccaro & Fleming, 2018) a identifié un réseau domaine-général impliquant le cortex préfrontal latéral et médian ainsi que l'insula et le précunéus, associé à la magnitude des jugements métacognitifs, que ce soit sur des tâches de mémoire ou de prise de décision. En prenant en compte le moment où les jugements étaient faits, cette méta-analyse a montré que les jugements rétrospectifs mettaient plus en jeu le cortex parahippocampique bilatéral et le gyrus frontal inférieur gauche alors que les jugements prédictifs activaient plus le cortex préfrontal médian, le cortex préfrontal dorsolatéral gauche et l'insula droite.

En résumé, toutes ces études montrent que malgré la généralité de certains facteurs ou indices, les processus spécifiques mis en jeu pour divers types de jugements diffèrent, notamment entre les jugements prédictifs et les jugements rétrospectifs. Les jugements rétrospectifs sont plus influencés par la force de la trace et la facilité de récupération en mémoire, alors que les prédictions, et en particulier les FOK, sont influencées par la familiarité avec l'indice et les informations partielles récupérées sur la cible à retrouver. Ainsi, il est donc important de considérer cette diversité et ces divergences lorsque l'on cherche à évaluer les capacités métacognitives de personnes. Dans le cadre de cette thèse, il était donc nécessaire d'utiliser plusieurs paradigmes (e.g., PoR, RCJ) afin d'avoir une vision plus étendue du rôle de la ME et de la MS dans les jugements métacognitifs puisque certaines conclusions obtenues avec un type de paradigme ne sont pas systématiquement transférables à un autre.

## 2.3. Métacognition épisodique et sémantique : vision dichotomique

La spécificité de la métacognition peut aussi être illustrée par des dissociations dans la capacité des personnes à évaluer leurs performances selon la forme de mémoire mis en jeu dans la tâche de premier ordre. Dans le domaine de la mémoire, les paradigmes de métacognition sur la mémoire déclarative les plus utilisés peuvent être divisés en deux classes selon qu'ils utilisent des tâches de premier ordre faisant appel à la mémoire épisodique ou sémantique (voir Mazancieux, Dinze, et al., 2020; Morson et al., 2015; Wojcik et al., 2013 pour des exemples d'études comparant les deux types tâches avec des jugements de FOK).

### 2.3.1. Mémoire épisodique et sémantique

Tulving (1972) a proposé que la mémoire déclarative, c'est-à-dire la mémoire à laquelle on a accès consciemment et que l'on peut explicitement rapporter, peut être divisée en deux types de mémoire : la ME et la MS (voir Renoult et al., 2019; Renoult & Rugg, 2020 pour des revues concernant cette distinction). La ME renvoie à la mémoire des événements vécus dans un contexte spatio-temporel spécifique. D'un autre côté, la MS renvoie aux connaissances générales qu'un individu possède sur le monde (Tulving, 1972 décrivait la MS en lien avec le langage mais, comme il en sera discuté dans le [Chapitre 3](#), cette conception a depuis évolué). Ces connaissances regroupent notamment les concepts, les liens entre ces concepts, les faits, et les règles. La MS peut être évaluée avec des questions de culture générale, des tâches de reconnaissance de personnes célèbres, ou des tâches de dénomination, qui testent les connaissances que les participants avaient déjà avant l'expérimentation.

Des processus communs sous-tendent les deux types de mémoire et la distinction entre les deux n'est pas stricte (Irish & Vatansever, 2020; Renoult et al., 2019; ce point sera rediscuté dans [Chapitre 3](#)). Cependant, certains processus et mécanismes seraient tout de même plus spécifiques à un type de mémoire qu'à l'autre. Par exemple, la récupération en ME s'accompagne d'une expérience de conscience auto-noétique (i.e., capacité à être conscient de soi dans le passé, le présent, et le futur) alors que la récupération en MS impliquerait un état de conscience noétique (i.e., conscience de savoir) (Tulving, 1985b). La conscience auto-noétique se caractérise par la sensation de voyage mental dans le temps et par le processus de recollection qui est la capacité à récupérer des détails contextuels spécifiques sur les épisodes vécus. La conscience noétique se manifeste quant à elle par une simple conscience de familiarité (i.e., impression d'avoir déjà rencontré le stimulus auparavant) ou de savoir (Tulving, 1985b).

Comme décrit lors de la description des paradigmes de jugements métacognitifs dans le [Chapitre 1](#) de ce travail, ces deux types de mémoire peuvent être évalués de manière explicite sur des tâches de rappel libre ou indicé dans lesquelles les participants doivent récupérer

l'information en mémoire d'eux-mêmes (comme c'est le cas dans de nombreux paradigmes de JOL), ou sur des tâches de reconnaissance (e.g., dans paradigme de FOK). Le rappel en ME repose sur la capacité à revivre mentalement un événement du passé et donc sur le mécanisme de recollection. D'un autre côté, plusieurs mécanismes peuvent être mis en jeu sur une tâche de reconnaissance épisodique. Dans une tâche de reconnaissance, il faut identifier quel stimulus a été traité dans une phase d'encodage parmi des distracteurs. Pour cela, une personne peut refaire l'expérience mentale de la perception du stimulus et se remémorer des éléments spécifiques au contexte d'encodage (i.e., recollection, spécifique aux tests de ME) ou estimer si le stimulus est familier (Yonelinas, 1994, 2002). La familiarité est un processus qui correspond à l'absence de recollection consciente et peut également être utilisée dans les tests de reconnaissance en MS (par définition, la familiarité désigne la sensation de connaître la réponse sans recollection).

En conclusion, la distinction entre ME et MS a largement influencé la recherche en psychologie cognitive et neuropsychologie de la mémoire. Cette distinction, longtemps ignorée dans la métamémoire, a plus tard également façonné une partie de ce champ. En effet, il est important de prendre en compte la forme de mémoire auquel la tâche fait appel lorsque l'on étudie les jugements métacognitifs sur des tâches de mémoire car des processus différents peuvent être mis en jeu. Ainsi, les demandes de la tâche de premier ordre peuvent différer selon le type de mémoire testé, ce qui pourrait également entraîner des divergences dans la façon d'évaluer les performances sur ces tâches. Nous appuyons sur le fait que nous ne supposons pas deux systèmes différents de mémoire, mais plutôt que nous considérons une dichotomie dans les tâches de premier ordre qui permettent d'évaluer le monitoring de la mémoire à long terme. Notamment, certaines tâches de mémoire font appel à des processus plus épisodiques (e.g., recollection) et d'autres plus sémantiques (e.g., connaissances, liens entre concepts).

### **2.3.2. Distinction épisodique et sémantique dans la métacognition**

Lorsque l'on s'intéresse à cette distinction épisodique/sémantique dans des tâches de métacognition, les résultats divergent. Assez peu d'études ont directement comparé les performances de participants sur des tâches de métacognition épisodiques et sémantiques (voir Mazancieux, Dinze, et al., 2020 pour une revue dans le FOK). Globalement, les travaux qui vont être présentés dans la partie suivante suggèrent que des processus communs sont mis en jeu dans les jugements rétrospectifs épisodiques et sémantiques. D'un autre côté, des processus semblent plus dépendants de la tâche de premier ordre (i.e., épisodique ou sémantique) dans les paradigmes dans lesquels les participants doivent prédire leur performance (e.g., FOK).

### 2.3.2.1. *Études corrélationnelles*

Une façon d'évaluer si des processus communs sous-tendent différents types de jugements est de corrélérer les performances métacognitives selon que la tâche de premier ordre porte sur la ME ou sur la MS. Une étude précédente conduite au sein de notre laboratoire (Mazancieux, Dinze, et al., 2020) a montré une spécificité des FOK et une généralité des RCJ. Dans cette étude, les participants réalisaient deux tâches de mémoire dans un ordre randomisé : une de ME et une de MS. Dans chacune de ces deux tâches, les participants devaient d'abord essayer de rappeler l'information cible avant de faire une prédiction quant à leur capacité à la reconnaître (FOK). Ensuite, les participants devaient sélectionner la bonne cible parmi deux propositions et ensuite indiquer à quel point ils étaient certains d'avoir choisi la bonne réponse (RCJ). Dans la tâche épisodique, les cibles correspondaient aux mots des paires apprises dans une phase d'encodage. Dans la partie sémantique, les cibles étaient les réponses à des questions de connaissances générales. L'efficacité métacognitive des eRCJ et des sRCJ était positivement corrélée ( $\rho = 0.47$  [0.13; 0.78]), ce qui n'était pas le cas pour les eFOK et sFOK ( $\rho = 0.22$  [-0.89; 0.99]). De plus, l'efficacité métacognitive était meilleure pour les eFOK que pour les sFOK alors qu'il n'y avait pas de différence significative entre l'efficacité métacognitive des eRCJ et des sRCJ. Ces résultats suggèrent donc une spécificité de l'efficacité métacognitive des FOK selon le type de mémoire sur laquelle la tâche de mémoire porte et une généralité de l'efficacité métacognitive des jugements rétrospectifs. Une explication possible proposée par les auteurs est que même si les FOK portent sur la réussite à la tâche de reconnaissance, les capacités de rappel sont utilisées comme indice. Ceci était notamment soutenu par la corrélation positive entre l'efficacité métacognitive et les performances de rappel, que ce soit pour les eFOK et pour les sFOK mais non pour les sRCJ ni les eRCJ. Or, comme discuté précédemment, la récupération en MS et le rappel en ME ne mettent pas en jeu les mêmes processus (savoir *versus* se remémorer des informations précises sur l'encodage) et n'impliquent pas le même état de conscience (noétique *versus* auto-noétique). D'un autre côté, les RCJ seraient influencés par des indices basés sur l'expérience subjective plus généraux (i.e., utilisables dans tous les domaines cognitifs) comme la fluence de récupération.

### 2.3.2.2. *Preuves neuropsychologiques de dissociations dans le FOK*

En outre, ce décalage de sensibilité métacognitive entre les eFOK et sFOK a été mis en évidence dans plusieurs populations cliniques et s'illustre par des dissociations d'altération/préservation. Par exemple, plusieurs études ont montré des patterns de préservation et d'altération de la sensibilité métacognitive dans la MA (voir Brandt et al., 2018; Ernst et al., 2015 pour des revues). En particulier, un déficit de sensibilité métacognitive des

eFOK a été retrouvé dans différents travaux (Cosentino et al., 2011, 2015; Souchay et al., 2002). Dans l'étude de Pappas et al. (1992), des personnes MA réalisaient une tâche de sFOK (répondre à des questions d'information générale) et une tâche de eFOK (retenir le dernier mot d'une phrase). Les participants ont rappelé moins de réponses que les contrôles que ce soit sur la tâche sémantique ou épisodique. Cependant, la sensibilité métacognitive des patients était plus faible que les contrôles sur le sFOK, bien que les performances en reconnaissance de premier ordre des personnes MA et contrôles ne soient pas significativement différentes. La sensibilité des patients MA était également altérée sur le eFOK. À l'inverse, d'autres études portant sur le sFOK n'ont pas montré de déficit de sensibilité dans cette pathologie (Bäckman & Lipinska, 1993; Lipinska & Bäckman, 1996). Dans ces études, même si les patients MA avaient moins bien répondu aux questions de culture générale que les contrôles, leur capacité à évaluer l'exactitude de leurs réponses semblait intacte. Ainsi, même s'il existe des divergences dans les résultats concernant le sFOK, ces études semblent indiquer un déficit en eFOK plus marqué que sur le sFOK dans la MA, du moins dans les stades précoces de la maladie. Ces observations ont amené Souchay (2007) à proposer l'hypothèse de fractionnement de la métacognition dans la MA. Ce fractionnement serait une conséquence du processus inférentiel des jugements métacognitifs : les indices mnémiques utilisés pour faire un FOK dépendraient de la tâche. Notamment, l'évaluation des connaissances générales reposerait sur des informations bien ancrées dans le système sémantique et en grande partie sur la familiarité avec la question alors que les jugements épisodiques reposeraient plus sur des processus auto-initiés comme la récupération partielle de la cible qui nécessiteraient la conscience auto-noétique.

Un pattern de dissociation a également été observé sur les tâches de FOK chez des patients avec des lésions frontales (Janowsky et al., 1989). Ces patients présentaient une bonne sensibilité métacognitive sur des tâches de sFOK. Cependant, les patients présentaient une altération de la sensibilité sur une tâche de eFOK, qui s'observait uniquement lorsqu'un délai était introduit entre l'apprentissage et le test (1 à 3 jours) et donc quand les capacités de rappel et la reconnaissance étaient diminuées. Ces résultats ont contribué à soutenir l'idée que les lobes frontaux sont davantage impliqués dans les eFOK que les sFOK.

D'autres populations cliniques comme chez les personnes avec un trouble du spectre autistique présentaient également un décalage eFOK/sFOK (TSA; Grainger et al., 2014; Wojcik et al., 2013; Wojcik & Souchay, 2013). Ces personnes avaient une moins bonne sensibilité métacognitive sur le eFOK que des personnes contrôles, mais une sensibilité équivalente aux contrôles sur le sFOK.

Enfin, plusieurs études ont mis en évidence une moindre sensibilité métacognitive sur les tâches de eFOK malgré des compétences préservées sur des tâches de sFOK chez les personnes âgées (e.g., Allen-Burge & Storandt, 2000; Isingrini et al., 2008; Marquié & Huet,

2000; Morson et al., 2015; Perrotin et al., 2006), bien que cet résultat n'ait pas toujours été observé (voir Eakin et al., 2014; Eakin & Hertzog, 2012). Plusieurs hypothèses ont été proposées pour expliquer ce déclin du eFOK : l'hypothèse de contrainte mnésique (*memory constraint hypothesis* ; Hertzog et al., 2010), l'hypothèse de déficit exécutif (Perrotin et al., 2006; Souchay et al., 2000), ou encore l'hypothèse de déficit de recollection (Souchay et al., 2007). Ces études et hypothèses seront expliquées plus en détail dans la revue et méta-analyse présentée dans le [Chapitre 4](#).

### 2.3.2.3. Études en neuroimagerie

En plus des études corrélationnelles et neuropsychologiques, des études en neuroimagerie suggèrent que les eFOK et sFOK reposent sur des substrats neuronaux distincts (e.g., Elman et al., 2012; Reggev et al., 2011). De précédents travaux en IRM fonctionnelle (IRMf) ont étudié séparément les bases du eFOK (Maril et al., 2003; Schnyer et al., 2005) et du sFOK (Kikyo et al., 2002; Kikyo & Miyashita, 2004). Par exemple, Maril et al. (2003) ont montré que l'activation du gyrus frontal inférieur gauche et dans le cortex pariétal gauche dépendait des ressentis subjectifs des participants sur une tâche épisodique : ils s'activaient plus lorsque les participants pensaient avoir correctement rappelé la cible que lorsqu'ils éprouvaient un FOK (i.e., ils pensaient pouvoir reconnaître la cible qu'ils n'arrivaient pas à rappeler). L'activation pour ce dernier cas de figure était également plus forte que lorsqu'ils ne pensaient pas pouvoir retrouver la cible du tout. L'activation du gyrus frontal inférieur était également positivement corrélée à la force du sFOK dans une autre étude (Kikyo et al., 2002). Mises en commun, ces études suggèrent que certaines régions comme le cortex frontal inférieur notamment, semblent impliquées à la fois dans le eFOK et le sFOK, mais d'autres aires ou réseaux semblent plus spécifiques. Cependant, des paradigmes différents étaient utilisés à travers ces études, limitant l'interprétation d'une distinction épisodique/sémantique.

À notre connaissance, seules deux études ont directement comparé les bases neurales du eFOK et du sFOK (Elman et al., 2012; Reggev et al., 2011). Des patterns d'activations différents entre le eFOK et le sFOK ont été mis en évidence. Dans l'étude de Reggev et al. (2011), tous les participants réalisaient une tâche de FOK qui comprenait des essais portant sur la ME et d'autres sur la MS. Dans une première phase, ils devaient créer une phrase en associant un proverbe et un nom (ville ou personne). Ils étaient ensuite testés sur leur mémoire des noms lorsque les proverbes étaient présentés (épisodique) ou sur des questions de culture générale (sémantique). Pour chaque essai, les participants devaient indiquer s'ils pouvaient rappeler la réponse (i.e., réponse « *know* »), s'ils ne pouvaient pas la rappeler mais pouvaient le retrouver parmi des options (i.e., FOK), ou s'ils ne pensaient pas pouvoir la retrouver du tout (i.e., réponse « *don't know* »). Des analyses de contrastes entre les deux types d'essais ont

révélé que certaines régions étaient communes aux deux types de jugements et d'autres plus spécifiques. Par exemple, le cortex préfrontal inférieur droit était plus activé pour les jugements « *don't know* » que les FOK dans les essais épisodiques mais plus activé pour les jugements FOK que « *don't know* » pour les essais sémantiques. Les auteurs suggèrent que ce pattern d'activation pourrait être relié à des stratégies de recherche en mémoire impliquées dans le monitoring. D'autres régions telles que le gyrus temporal médian gauche, le cunéus, le cortex cingulaire postérieur étaient également marqueurs de FOK dans la tâche épisodique mais pas dans la tâche sémantique.

Dans l'étude d'Elman et al. (2012), les participants réalisaient également une tâche de FOK alternant des blocs d'essais faisant appel à la ME et d'autres portant sur la MS et devaient donner un FOK sur une échelle en 4 points. Contrairement aux tâches classiques de FOK, les participants ne devaient pas essayer de rappeler la cible avant de faire leur prédiction de reconnaissance. Certaines régions comme le cortex préfrontal ventro-latéral et le gyrus cingulaire antérieur étaient activées à la fois pour les eFOK et les sFOK faibles. De plus, d'autres régions étaient plus activées pour les eFOK élevés que pour les sFOK élevés (e.g., cortex pariétal postérieur ventral gauche, précunéus, cortex frontal). À l'inverse, d'autres régions notamment impliquées dans le traitement sémantique étaient plus activées dans les sFOK élevés que dans les eFOK élevés (e.g., lobe temporal antérieur droit).

## 2.4. Conclusion du Chapitre 2

En conclusion, dans ce chapitre, nous avons montré que les différents jugements de monitoring sont inférés sur la base d'indices qui peuvent être basés sur l'expérience subjective ou des croyances/théories naïves que les personnes possèdent. Ces indices peuvent notamment être de nature sémantique (e.g., fluence conceptuelle, informations partielles). De plus, selon le type de jugement, différents indices sont utilisés. Des dissociations neuropsychologiques et neuro-fonctionnelles existent entre les jugements rétrospectifs et les jugements prospectifs, mais également entre les jugements prédictifs portés sur la ME et la MS.

Le but de ce travail de thèse est double. Dans un premier temps, nous avons cherché à approfondir la compréhension des disparités entre métacognition épisodique et métacognition sémantique dans certaines populations (Chapitre 4 et 5). Un deuxième objectif était aussi de compléter les connaissances sur l'influence de certains indices sur les jugements métacognitifs : ceux qui reposent sur le traitement sémantique des informations sur les tâches de ME (Chapitre 6 et 7).

## CHAPITRE 2 : POINTS IMPORTANTS

- Les jugements métacognitifs sont inférentiels : ils sont basés sur des **indices**.
- Malgré des points communs, tous les jugements métacognitifs ne sont pas égaux : prédictif/ rétrospectif, épisodique/ sémantique. Il est important de prendre en compte ces divergences.
- Les RCJ semblent être plus **domaine-généraux** alors que les FOK seraient plus **mémoire-spécifiques**.
- Dans cette thèse, la **distinction épisodique/sémantique** dans la métacognition a été étudiée (Chapitres 4 et 5).

# Chapitre 3. Lien entre premier ordre et second ordre

---

Dans les études de la métacognition, les personnes jugent leur performance sur une tâche cognitive de premier ordre. Lorsque l'on veut évaluer l'exactitude des jugements métacognitifs, on compare l'évaluation subjective faite par les participants à leur performance réelle sur la tâche de premier ordre. Comme nous en avons discuté dans le [Chapitre 2](#), les performances métacognitives peuvent diverger selon le type de tâche de premier ordre présenté et le moment où le jugement métacognitif est fait. De plus, il est important de considérer la tâche de premier ordre et les processus cognitifs communs à la tâche de premier ordre et aux performances de second-ordre. Dans ce chapitre, le but est de montrer que sur des tâches de métamémoire, les capacités et processus mnésiques sont importants à considérer lorsque l'on cherche à étudier les compétences métacognitives. Plus particulièrement, la mémoire a un rôle dans la métacognition puisque certaines informations mnésiques épisodiques et sémantiques sont utilisées pour faire des jugements métacognitifs.

## 3.1. Lien statistique et conceptuel

Conceptuellement parlant, la sensibilité métacognitive est une habilité différente des capacités mnésiques. En théorie, évaluer si une personne est capable de juger l'exactitude de ses souvenirs et ses connaissances est différent d'évaluer si cette personne arrive à bien se souvenir, et ces deux capacités devraient être indépendantes (Nelson, 1984). Cependant, en pratique, les compétences métacognitives sont souvent reliées aux compétences mnésiques. Notamment, les modèles de métacognition supposent une relation entre les décisions métacognitives (i.e., de second ordre) et la tâche de premier ordre sur lesquelles ces décisions portent, d'un point de vue statistique et d'un point de vue plus cognitif.

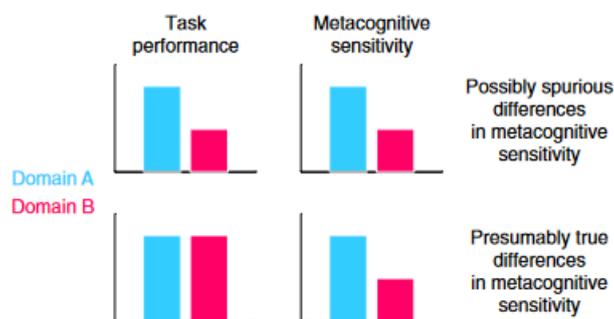
### 3.1.1. Dépendance au niveau de la mesure ?

Tout d'abord, la question de la relation entre la mémoire et la métamémoire se pose d'un point de vue statistique, vis-à-vis des mesures utilisées pour évaluer la métacognition. En particulier, lorsque l'on veut estimer les compétences métacognitives, il est important de prendre des précautions au niveau de la difficulté de la tâche de premier ordre avant de tirer des conclusions. En effet, même si certaines mesures de sensibilité métacognitive comme les corrélations gamma ont été développées pour être indépendantes de la performance de premier ordre, des travaux ont montré que ce n'était pas nécessairement le cas (Vuorre &

Metcalfe, 2021). Vuorre & Metcalfe (2021) ont par exemple observé une corrélation positive entre les performances sur une tâche de mathématique et la sensibilité métacognitive des jugements rétrospectifs sur cette même tâche mesurée par des corrélations gamma et des AUROC2. De plus, d'autres études ont mis en évidence que la sensibilité métacognitive était confondue avec la performance de type-1. Par exemple, il serait plus facile d'évaluer sa performance sur une tâche de premier ordre facile plutôt que difficile (Galvin et al., 2003; Masson & Rotello, 2009). Ces observations ont notamment amené le développement de nouvelles méthodes pour essayer de mesurer la sensibilité métacognitive indépendamment de la performance de type-1 (e.g., efficacité métacognitive, M-ratio ; Fleming et al., 2016; Maniscalco & Lau, 2012). Cependant, il n'est pas pertinent d'utiliser ces mesures dans tous les cas de figure car elles requièrent certaines contraintes dans les paradigmes (e.g., grand nombre d'essais, tâche de reconnaissance avec seulement deux alternatives) qui ne sont pas toujours possibles ou pertinentes à mettre en place.

Une récente méta-analyse (Rouy et al., 2021) a montré qu'un déficit de sensibilité métacognitive chez des personnes présentant un trouble schizophrénique était réduit (voire disparaissait) lorsque seules les études ayant contrôlé la performance à la tâche de premier ordre étaient considérées. Cette méta-analyse incluait des études dont les tâches de premier ordre étaient des tâches de mémoire, de perception, de cognition sociale et d'agentivité. De plus, parmi les études n'ayant pas contrôlé pour la performance cognitive, une corrélation positive était observée entre la performance de premier ordre et la sensibilité métacognitive. Ces résultats corroborent l'hypothèse selon laquelle le déficit cognitif observé chez certaines personnes, s'il est mal contrôlé, peut amener à une surestimation du déficit métacognitif.

Ces précautions vis-à-vis du contrôle de la tâche de premier ordre sont à prendre en compte, à la fois lorsque les compétences métacognitives sont comparées dans plusieurs domaines cognitifs, comme c'est le cas sur la Figure 3.1, mais peuvent également s'étendre aux études dans lesquelles plusieurs conditions sont comparées sur une même tâche.



**Figure 3.1** : Exemples d'interprétations sur les différences entre performance de premier ordre et sensibilité métacognitive (figure issue de Rouault et al., 2018).

Ainsi, ces travaux montrent que les performances de second ordre et de premier ordre sont reliées. Cela implique que la performance à la tâche de premier ordre est importante à prendre en compte lorsque l'on veut tirer des conclusions sur les capacités métacognitives. Notamment, lorsque l'on compare ces capacités sur plusieurs tâches ou à travers plusieurs conditions expérimentales, il faut essayer de minimiser les différences de performances de type-1 pour évaluer une potentielle évolution de la sensibilité métacognitive « pure ».

### **3.1.2. Dépendance au niveau cognitif ?**

Les performances mnésiques et métacognitives peuvent être liées au-delà de possibles artefacts liés à la mesure de la sensibilité métacognitive. En effet, les processus de mémoire, tels que la remémoration, peuvent être directement utilisés comme indice pour les jugements, en particulier pour les jugements prédictifs (Isingrini et al., 2016; Souchay et al., 2007). Cela implique que si les jugements métacognitifs se basent sur ces processus de mémoire qui sont également mis en jeu dans la tâche de premier ordre, et si ces processus sont altérés ou réduits, les performances métacognitives peuvent en être impactées. Comme présenté dans le [Chapitre 2](#) (p. 30), certaines personnes présentent une diminution voire un déficit de sensibilité métacognitive sur des tâches de FOK lorsque la tâche de premier ordre met en jeu la ME mais non lorsqu'elle porte sur la MS (e.g., personnes âgées, malades d'Alzheimer, trouble du spectre autistique). Or, ces populations présentent également des particularités de la ME mais des compétences sémantiques relativement préservées voire élevées. Ces observations contribuent donc au questionnement sur la relation entre compétences mnésiques et compétences métacognitives.

Tout d'abord, une précédente étude (Mazancieux, Dinze, et al., 2020) a mis en évidence une corrélation positive entre la proportion de rappel et l'efficacité métacognitive. Cette corrélation était observée sur les tâches de sFOK et de eFOK, mais non sur les tâches de sRCJ ou eRCJ. Ce résultat montre une relation spécifique entre performance mnésiques et sensibilité métacognitive dans des paradigmes de FOK mais non pour les jugements rétrospectifs. Une explication est que les FOK sont généralement faits après une tentative de rappel. Notamment, lorsqu'elles doivent prédire leur capacité à reconnaître la cible (i.e., PoR), les personnes peuvent s'engager dans la recherche de la cible et utiliser le résultat de cette recherche pour prédire leur performance future (Koriat, 1993; Koriat et al., 2003; Thomas et al., 2012). Cette recherche en mémoire met en jeu le processus de recollection puisque les personnes doivent se rappeler ce qu'elles ont appris dans le contexte d'une phase d'apprentissage. Ainsi, même dans les cas où la cible n'est pas rappelée, les informations partielles accessibles au moment de faire une prédiction de reconnaissance peuvent être de nature épisodique (e.g., « je me souviens que cette paire m'avait fait penser à mon chat »).

Pour soutenir cette idée, plusieurs études ont montré une relation entre jugements métacognitifs et mesures de recollection. Par exemple, de précédents travaux ont montré un lien entre recollection/familiarité et FOK (Hicks & Marsh, 2002; MacLaverly & Hertzog, 2009). Dans l'expérience 1 de Hicks & Marsh (2002), les participants réalisaient une tâche de FOK classique (i.e., paradigme rappel-jugement-reconnaissance), et pour chaque réponse sur la tâche de reconnaissance finale, ils indiquaient s'ils avaient répondu au hasard (« *guess* », G), s'ils s'étaient basés sur de la familiarité (« *know* », K) ou s'ils se rappelaient de détails précis (« *remember* », R). Les FOK étaient positivement corrélés à l'expérience associée à leur réponse future. En effet, les essais pour lesquels les participants avaient donné des FOK élevés étaient ceux qui étaient par la suite associés à des jugements R. Les essais qui avaient reçus des FOK médiums étaient plus tard plus associés à des réponses R et K. Enfin, les essais ayant reçu de FOK faibles étaient autant associés à des réponses R, K que G. De plus, la corrélation entre FOK et jugements R/K sur la tâche de reconnaissance était également retrouvée dans l'étude de MacLaverly & Hertzog (2009). Ces résultats montrent que les FOK reflètent l'expérience subjective future associée à la reconnaissance. Encore une fois, ces exemples laissent supposer que des mécanismes mnésiques guident les jugements métacognitifs.

En outre, d'autres études ont mis en évidence l'utilisation du processus de recollection pour les FOK. En particulier, le processus de recollection contribuerait à la sensibilité des FOK (Isingrini et al., 2016; Souchay et al., 2007). Par exemple, Isingrini et al. (2016) ont utilisé un paradigme classique de eFOK dans lequel ils ont ajouté une phase de jugement Remember/Know/New. Lorsque les participants n'arrivaient pas à rappeler la cible, ils faisaient un FOK puis ils devaient indiquer le degré d'*awareness* associé à chaque essai. Ils indiquaient « R » s'ils se remémoraient de quelconques informations concernant l'apprentissage de la paire (i.e., recollection), « K » s'ils se souvenaient avoir vu l'indice mais qu'ils n'avaient aucun souvenir additionnel relatif à l'expérience d'apprentissage (i.e., familiarité), et « N » s'ils ne se souvenaient pas avoir vu l'indice. La sensibilité métacognitive était plus élevée pour les essais associés à de la recollection que pour les essais associés à des réponses K (i.e., souvenir sans recollection, familiarité), et que pour les essais pour lesquels les participants n'avaient aucun souvenir. Lorsque les participants n'avaient aucun souvenir de l'indice, la sensibilité métacognitive était au niveau du hasard. Ces résultats montrent que les processus mnésiques de recollection et de familiarité sont nécessaires pour une bonne sensibilité des eFOK, et que la recollection est plus bénéfique que la familiarité.

De plus, selon l'hypothèse de recollection non-critérielle (G. A. Brewer et al., 2010), avoir accès à des informations contextuelles concernant les conditions d'encodage des stimuli, même si erronées, fourniraient des indications sur la trace mnésique. Ainsi, même si elles ne sont pas pertinentes pour le rappel de la cible (d'où le nom « non-critérielles »), ces

informations influenceraient les prédictions. Dans une étude (Schwartz et al., 2014), les participants ont réalisé une tâche de eFOK. Les stimuli à apprendre étaient des dessins d'animaux imaginaires associés à un nom inventé qu'ils devaient mémoriser. Plusieurs conditions étaient créées en fonction de la quantité d'informations contextuelles fournies au moment de l'encodage. Dans une première condition (pauvre), on ne leur donnait que le nom de l'animal. Dans une seconde condition (intermédiaire), son nom ainsi que son pays d'origine étaient fournis. Enfin, dans une troisième condition (riche), le nom, le pays, le régime alimentaire ainsi que le poids de l'animal étaient donnés. Dans la phase de test, lorsqu'ils n'arrivaient pas à rappeler le nom de l'animal présenté, les participants devaient essayer de rappeler les informations supplémentaires apprises sur l'animal. Ils prédisaient ensuite leur capacité à reconnaître le bon nom (i.e., FOK) puis à la fin, ils devaient reconnaître le bon nom parmi 4 propositions. La magnitude des FOK était plus élevée dans les cas où les participants se rappelaient d'informations contextuelles par rapport aux essais dans lesquels ils ne s'en rappelaient plus. Cependant, la quantité d'informations fournies n'influçait pas la sensibilité des prédictions. Cette étude montre que, dans une certaine mesure, les informations récupérées en ME influencent la magnitude des jugements métacognitifs, même lorsqu'elles ne concernent pas directement la cible à récupérer, bien que cela n'influence pas nécessairement la sensibilité des jugements.

Les études précédemment décrites montrent que l'intervention des processus de ME sur des tâches de prédictions de reconnaissance, et cela pourrait expliquer, du moins en partie, les déficits métacognitifs observés sur les tâches de eFOK dans certaines populations (e.g., Hertzog et al., 2010; Souchay, 2007; Souchay et al., 2007). En effet, dans les études réalisées chez les personnes ayant des troubles de la mémoire comme dans la MA, certains auteurs proposent que le déficit de sensibilité métacognitive observé sur certaines tâches épisodiques serait la conséquence du déficit de ME (e.g., Souchay et al., 2002 ; voir Ernst et al., 2015 pour une revue). Plus particulièrement, Souchay (2007) a proposé que le manque de conscience autoéotique au moment de la récupération en mémoire était à l'origine du déficit de sensibilité métacognitive spécifique au eFOK dans la MA. Ainsi, les personnes auraient moins accès aux indices épisodiques comme les informations contextuelles, ce qui pourrait engendrer une moindre sensibilité. Les personnes MA se reposeraient plus sur le processus de familiarité pour faire leurs jugements, car ce processus est préservé dans cette pathologie, ce qui aiderait moins à avoir une sensibilité optimale que la recollection (Isingrini et al., 2016).

De plus, certains auteurs proposent que la diminution de la sensibilité des eFOK parfois observée dans le vieillissement sain, bien que pas toujours retrouvée (e.g., MacLaverly & Hertzog, 2009), découlerait également du déclin en ME (Hertzog et al., 2010; Souchay et al.,

2007). Ces hypothèses ainsi que des hypothèses alternatives (e.g., hypothèse exécutive, Perrotin et al., 2008) seront discutées dans le [Chapitre 4](#) de ce travail de thèse.

Cependant, il est important de noter que ce lien entre performance cognitive et sensibilité métacognitive n'est pas toujours retrouvé. En effet, un déficit en ME n'entraîne pas systématiquement un déficit métacognitif, même sur des tâches utilisant des jugements généralement influencés par les capacités mnésiques tels que le FOK. Par exemple, plusieurs études ayant été menées auprès de patients amnésiques (i.e., avec un trouble de la ME) présentant une atrophie des lobes temporaux mais sans atrophie des lobes frontaux n'ont pas mis en évidence de déficit spécifique de sensibilité métacognitive chez ces personnes. En effet, que ça soit pour le eFOK ou le sFOK, les patients étaient aussi aptes que le groupe contrôle à évaluer leur performance future, même lorsque leur performance de type-1 était moindre que celles des contrôles (Janowsky et al., 1989; Shimamura & Squire, 1986). De même, cette dissociation entre capacités mnésiques et habiletés métacognitives était observée chez des patients avec épilepsies temporales (Howard et al., 2010). Ces patients présentaient des performances altérées sur une tâche de ME, en rappel et en reconnaissance, mais les analyses ne révélaient pas de différence de sensibilité de FOK sur la même tâche par rapport à des participants contrôles. Cette dissociation entre altération mnésique et préservation métacognitive a également été observée dans d'autres populations comme le trouble schizophrénique où l'altération de la MS (rappel) et de la ME (reconnaissance) ne s'accompagnait pas d'un trouble de sensibilité métacognitive de sFOK ni de eFOK (Bacon et al., 2001; Bacon & Izaute, 2009).

Par ailleurs, dans de nombreuses études réalisées auprès de participants sains, des modulations de performance de premier ordre n'étaient pas systématiquement accompagnées de différences sur les mesures de métacognition. Par exemple, dans une étude (Schacter, 1983), des paires de mots étaient apprises sous différentes conditions d'encodage (temps de présentation court ou long) et de rétention (test immédiat ou 1 semaine plus tard). Les manipulations d'encodage et de rétention affectaient la performance mnésique puisque les participants rappelaient et reconnaissaient moins de cibles présentées pendant moins longtemps à l'encodage plutôt que celles présentées plus longtemps, et également moins après 1 semaine qu'immédiatement après avoir appris les paires. Cependant, la performance métacognitive, mesurée ici par la proportion d'essais dont la cible a été reconnue selon qu'ils aient reçu un FOK positif (i.e., les participants pensaient reconnaître la cible) ou un FOK négatif (i.e., les participants ne pensaient pas reconnaître la cible), n'était pas affectée par ces manipulations. De même, une autre étude a montré que la consommation aiguë d'alcool pouvait altérer la capacité à répondre à questions de MS, sans pour autant affecter la sensibilité des sFOK (Nelson et al., 1986). Des résultats similaires ont été obtenus par l'injection d'autres types de drogues (e.g., Merritt et al., 2005 ; mais voir Bacon et al., 1998).

En conclusion, même si de nombreuses études suggèrent un lien entre premier-ordre et second-ordre, ce lien est complexe. Alors que certains processus semblent communs entre la tâche de premier ordre et la tâche de second ordre (e.g., recollection dans la ME et le eFOK), d'autres sont plus spécifiques. Les performances métacognitives sont en partie reliées aux performances cognitives car les mesures de métacognition peuvent être affectées par les performances à la tâche de premier ordre, mais également parce que les processus de mémoire (e.g., épisodiques) peuvent être utilisés pour faire des jugements métacognitifs, en particulier dans les jugements prédictifs. Cependant, une altération de la performance sur la tâche de premier ordre n'entraîne pas systématiquement une altération de la sensibilité métacognitive. Il semble donc important d'approfondir les connaissances sur les capacités métacognitives de personnes présentant des pathologies de la mémoire. Si le lien entre capacités épisodiques et compétences métacognitives a été investigué dans de nombreux travaux, peu d'études ont été portées sur l'importance de processus sémantiques. Notamment, est-ce qu'un déficit sémantique peut entraîner un déficit métacognitif sur certaines tâches ? Comment ces processus sémantiques interviennent dans les jugements métacognitifs ?

### **3.2. Traitement sémantique : utilisation dans la mémoire épisodique et la métacognition épisodique**

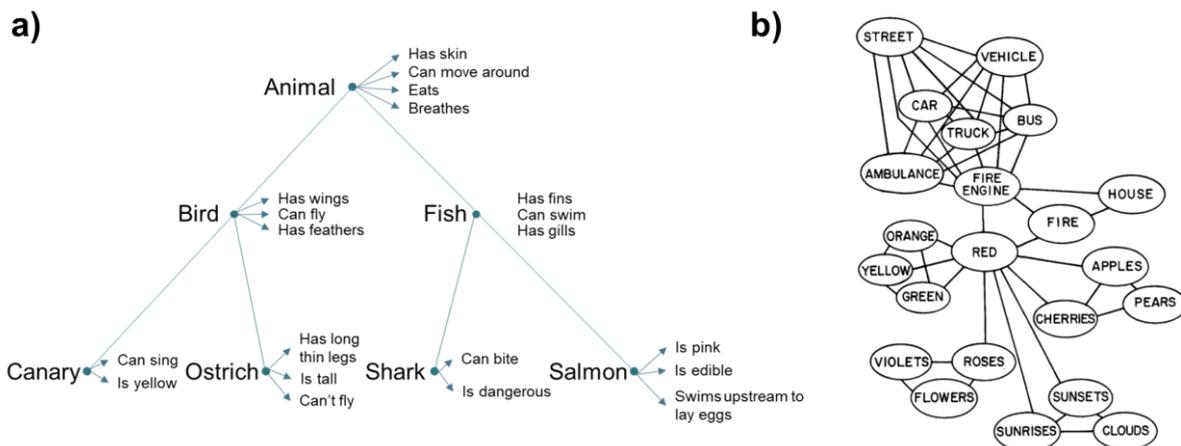
Bien que la ME et la MS soient souvent étudiés en comparaison ou opposition, leur interaction serait primordiale pour un fonctionnement mnésique efficace (Renoult et al., 2019; Renoult & Rugg, 2020; Tulving, 1983). Notamment, l'importance des traitements sémantiques pour l'encodage et la récupération des informations épisodiques a été très documentée. De plus, comme cela a été abordé dans le [Chapitre 2](#), certaines informations sémantiques peuvent être utilisées pour les jugements métacognitifs, même lorsque la tâche de premier ordre est une tâche de ME.

#### **3.2.1. Traitement sémantique : définition et modèles**

La MS regroupe les connaissances générales et les croyances sur le monde et sur le soi, et les représentations multimodales (ou amodales selon les auteurs) de concepts sans référence à leur contexte spatial et temporel d'apprentissage. Si la MS était à l'origine principalement présentée comme la mémoire permettant le langage et le stockage des symboles verbaux, des mots et de leurs référents, et très reliée à la compréhension du langage (Tulving, 1972), sa conception a grandement évolué au fil du temps. En effet, l'intervention de la MS dans les processus langagiers a été démontrée (e.g., Federmeier et al., 2002; Kutas & Federmeier, 2000; Levelt, 1992) et reprise dans de nombreux modèles du langage (e.g.,

Levelt, 1992, 2001). De plus, le langage joue un rôle important dans la représentation des informations en MS. Cependant, le langage n'est pas considéré comme nécessaire à la formation et à la récupération d'informations sémantiques (Tulving, 1983).

Plusieurs modèles ont été tenté de décrire l'organisation de la MS (e.g., Binder & Desai, 2011; Chen et al., 2017; Chen & Rogers, 2014; Collins & Loftus, 1975; Patterson & Lambon Ralph, 2016 ; voir Yee, 2019; Yee et al., 2014, 2018 pour des revues). Initialement, des modèles cognitivistes proposaient que la MS était organisée en niveaux selon une hiérarchie (e.g., Collins & Loftus, 1975; Collins & Quillian, 1969, voir Figure 3.2). Ces modèles suggéraient que chaque concept était représenté par un nœud dans le réseau et relié à d'autres concepts par des relations abstraites (e.g., des règles, des procédures). Par exemple, on peut savoir qu'un cheval respire car le concept de cheval appartient à la catégorie des animaux dont une des caractéristiques est qu'ils respirent. Ces modèles supposaient qu'il était possible d'accéder à des caractéristiques d'un concept par une propagation de l'activation vers d'autres concepts. Ces modèles ont cependant peu de bases neuroanatomiques.



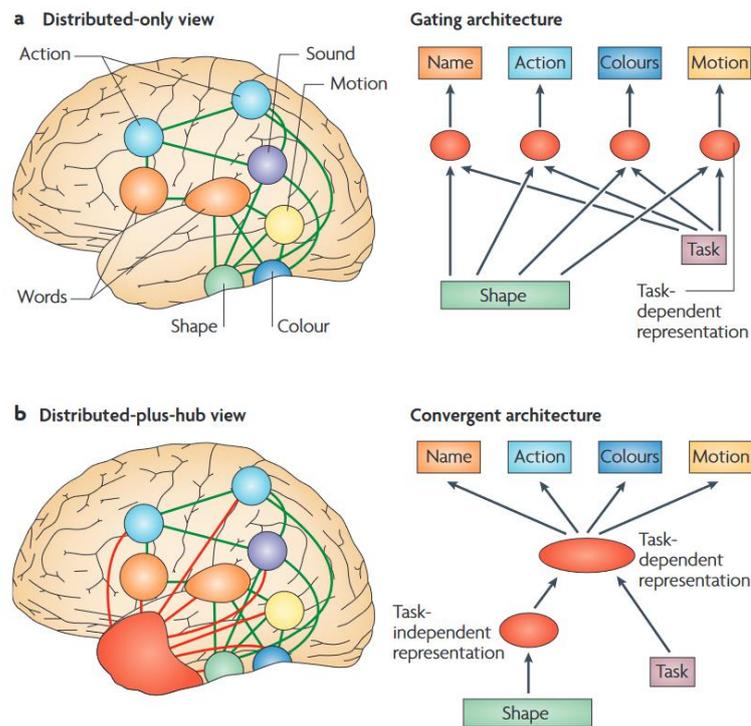
**Figure 3.2 : Exemples de modèles cognitivistes.**

- a) Modèle hiérarchique reproduit de Collins & Quillian (1969). Chaque trait et flèche représente une relation.
- b) Modèle distribué issu de Collins & Loftus (1975). Chaque bulle représente un concept. La longueur des traits représente la force d'association entre deux concepts.

Des théories plus récentes et neuro-inspirées, comme l'hypothèse *hub-and-spoke* (Lambon Ralph, 2014; Patterson et al., 2007; Patterson & Lambon Ralph, 2016 ; Figure 3.3b) supposent que des régions modalités-spécifiques supportent chaque caractéristique d'un concept (e.g., couleur, forme, nom, etc...). Cette hypothèse reprend notamment l'idée développée par des théories de la cognition incarnée qui supposent que la MS fait partie d'un système de mémoire intégré dans les systèmes sensori-moteurs et distribué à travers

plusieurs régions cérébrales (e.g., Barsalou, 1999; Martin, 2007; Pulvermüller, 2013). Selon les théories incarnées, et plus particulièrement les théories « uniquement distribuées » (*distributed-only*; Figure 3.3a), les concepts sont formés par des liens directs entre les systèmes sensori-moteurs et par l'activation simultanée des régions modalité-spécifiques. Ainsi, on accéderait aux caractéristiques d'un concept (e.g., forme d'un objet) par une simulation des systèmes sensori-moteur activés lors de l'expérience de l'objet réel dans le monde (i.e., régions activées pendant la perception et l'action). Par exemple, la connaissance que l'on a du concept de cheval regroupe les propriétés visuelles (e.g., taille, forme), mais aussi ses propriétés de mouvement (e.g., caractéristiques du galop), l'odeur associée, la sensation tactile, etc... chacune de ces propriétés étant représentée au niveau des régions qui traitent chaque type d'information lors de la perception d'un cheval dans le monde. L'activation d'une région spécifique peut se diffuser et activer les autres régions. Ainsi, si la tâche est de dénommer une image, voir la forme du concept (e.g., dessin d'un cheval) activerait les régions impliquées dans la représentation des formes puis cette activation se diffuserait auprès des régions sous-tendant les représentations langagières. L'hypothèse *hub-and-spoke* suppose qu'en plus des liens directs entre elles, ces régions communiquent également par des liens appelés *spokes* (traits rouges sur la Figure 3.3b), avec un hub amodal (zone rouge sur Figure 3.3b) situé dans le lobe temporal antérieur. Ce hub représente l'abstraction des propriétés spécifiques et les associations entre ces propriétés, et constitue donc une zone de convergence. À l'inverse des théories uniquement distribuées, cette hypothèse suppose que peu importe la tâche, le même réseau neuronal est activé (i.e., hub) et que les régions modalité-spécifiques communiquent également à ce niveau. Ainsi, selon cette hypothèse, le hub serait activé peu importe la modalité d'entrée (e.g., qu'il faille dénommer une image ou visualiser mentalement l'image correspondant à un mot, mais voir par exemple Gainotti, 2017 pour une remise en question).

La MS permet de créer du sens et d'organiser le monde grâce à des mécanismes d'abstraction de régularités à travers plusieurs épisodes (rejoint la notion de schémas, Bartlett, 1932), et de généralisation, c'est-à-dire l'application des connaissances apprises à d'autres contextes (Yee, 2019). La majorité du contenu de la MS se construit à travers ce que l'on perçoit et nos actions. Ce type de mémoire nous permet ainsi de faire des prédictions sur le monde sur la base de ce qui a été appris ou vécu auparavant. De nombreuses activités de la vie de tous les jours dépendent de la MS. Notamment, la MS est utile pour reconnaître, dénommer et utiliser des objets (Hodges, 2000), identifier des personnes, savoir à quoi un mot fait référence, mais également pour créer des associations entre des concepts (Bozeat et al., 2000). La capacité à représenter des concepts sous forme abstraite permet ainsi aux humains de communiquer, de comprendre le monde et d'agir dessus.



**Figure 3.3** : Théories neuro-inspirées de l'organisation du système sémantique. Schémas issus de Patterson et al. (2007).

Les théories uniquement distribuées supposent que les connaissances conceptuelles se forment par des liens directs (en vert) entre les multiples régions modalités-spécifiques. La théorie hub-and-spoke suppose qu'en plus de ces liens directs, les associations entre ces régions sont représentés dans un hub situé dans le lobe temporal antérieur (en rouge).

La MS comprend les représentations des concepts telles que décrites précédemment mais également les mécanismes de contrôle cognitif qui permettent d'y accéder et de les manipuler plutôt sous-tendus par les régions frontales (Chapman et al., 2020; Gotts & Plaut, 2002). Une altération de ce système peut entraîner, comme c'est le cas dans la démence sémantique (DS, Hodges et al., 1992; Snowden et al., 1989; Warrington, 1975) ou l'aphasie sémantique résultant d'un AVC, un trouble sur des tâches impliquant la MS (e.g., tâche de dénomination, d'identification, d'association sémantique). La DS entraînerait une perte des représentations sémantiques alors que l'aphasie sémantique se manifesterait plus par des difficultés à accéder aux représentations stockées (Corbett et al., 2009; Jefferies & Lambon Ralph, 2006; Warrington & Cipolotti, 1996; Warrington & McCarthy, 1983 ; mais voir Chapman et al., 2020).

Dans ce travail, nous appellerons traitement sémantique tout traitement faisant intervenir les représentations sémantiques. Plusieurs méthodes peuvent être utilisées pour manipuler l'importance et l'influence des traitements sémantiques sur la mémoire. Par exemple, on peut utiliser des mots ou des pseudo-mots (i.e., mots pouvant être lus mais qui n'ont pas de sens)

pour évaluer l'importance d'avoir une représentation sémantique stockée (e.g., Gardiner & Java, 1990; Hulme et al., 1991). D'un autre côté, comparer le traitement des mots concrets et des mots abstraits fait varier la richesse des informations reliées aux concepts (Paivio, 1990, 2010 ; e.g., Begg et al., 1989; Hertzog et al., 2003, 2014; Tauber & Rhodes, 2012). On peut également faire varier le degré d'association entre deux ou plusieurs mots, ou entre des images, pour manipuler la force des liens préétablis entre des concepts (e.g., Epstein et al., 1975; Gutchess & Park, 2009; Naveh-Benjamin et al., 2003), ou manipuler la fréquence des stimuli pour permettre d'évaluer l'effet de la robustesse des représentations (e.g., Begg et al., 1989; Gardiner & Java, 1990).

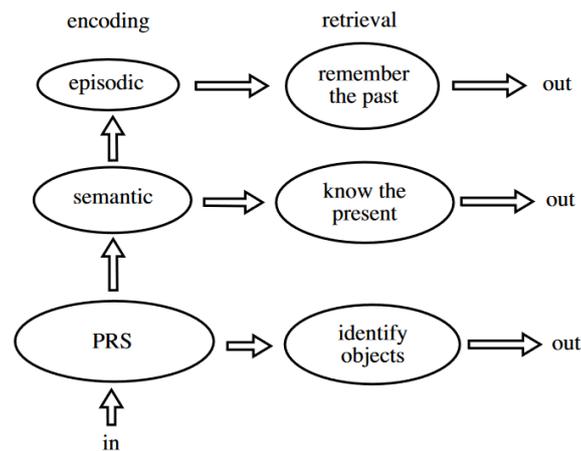
### **3.2.2. Utilisation de la mémoire sémantique dans la mémoire épisodique**

Certains travaux suggèrent que même si la ME et la MS reposent en partie sur des régions cérébrales et processus distincts (comme présenté dans le [Chapitre 2](#)), les deux types de mémoire partagent également des caractéristiques communes et interagissent. Plusieurs questions entourent le débat sur la distinction de la ME et de la MS. Notamment, la distinction pure entre les deux types de mémoire a été remise en question et leurs points communs ainsi que leur interaction est un sujet d'actualité (voir Irish, 2020; Irish & Vatansever, 2020; Renoult et al., 2019). Deux aspects de cette interaction sont entre autres discutés dans la littérature actuelle. Dans un premier temps, certains cherchent à comprendre à quel point la MS dépend de la ME pour se former. Ce débat a émergé de l'idée que la MS se formerait à travers l'extraction de régularités communes à plusieurs épisodes. D'un autre côté, d'autres recherches tendent à comprendre dans quelle mesure la mémoire sémantique influence et est primordiale pour la construction de la ME. C'est sur ce deuxième point que nous allons nous attarder.

Les connaissances sémantiques sont utilisées pour interagir avec le monde qui nous entoure et pour encoder et récupérer des événements épisodiques ainsi que le contexte spatio-temporel dans lequel ceux-ci ont été vécus. Lorsque l'on fait l'expérience d'un épisode, il y a une recombinaison des éléments contenus dans l'épisode, sur lesquels on possède de riches connaissances. Par exemple, pour se rappeler que l'on est allé en vacances à Florence, on a d'abord dû traiter les concepts de ville et de voyage ou tout autre élément perçu (e.g., ponts, tableaux, bâtiments), mais également ce qui concerne cette ville (e.g., son nom, sa localisation, la langue parlée).

L'idée de l'importance de la MS sémantique pour la ME n'est pas nouvelle. Par exemple, dans son modèle SPI (sériel, parallèle, et indépendant, voir Figure 3.4), Tulving (1995) supposait trois systèmes de mémoire, organisés hiérarchiquement du plus bas au plus haut niveau : le système perceptif, le système sémantique et le système épisodique. Selon ce

modèle, les caractéristiques physiques des objets seraient d'abord traitées par le système perceptif qui rendrait ces informations disponibles aux autres systèmes. L'information pourrait ensuite passer au niveau supérieur (i.e., sémantique), puis, lorsque l'information est traitée par le système sémantique, elle peut être stockée à ce niveau et/ou elle peut passer au niveau encore supérieur (i.e., épisodique). Selon ce modèle, l'encodage en mémoire est sériel (i.e., l'information doit d'abord être traitée par les niveaux inférieurs pour passer aux niveaux supérieurs), le stockage est parallèle (i.e., les informations de différentes natures sont stockées dans différents systèmes) et la récupération est indépendante (i.e., une fois qu'une information est passée dans les niveaux supérieurs, elle ne dépend plus des niveaux inférieurs). Ainsi, ce modèle suggère qu'une information est d'abord traitée par le système perceptif puis passe par le système sémantique afin d'être traitée et encodée dans le système épisodique. D'après ce modèle l'encodage en ME reposerait donc sur la préservation du système de MS.



**Figure 3.4 :** Modèle SPI (issu de Tulving, 1995).

Les 3 systèmes sont représentés dans la colonne de gauche. Les flèches indiquent le sens de circulation des informations encodées. Ce modèle a depuis été remis en question.

Certaines études ont cependant montré que supposer une hiérarchie n'était pas nécessaire. Notamment, des études conduites dans la DS ont remis en question ce modèle puisqu'elles ont montré des capacités d'apprentissage et de récupération en ME préservées malgré une altération du traitement sémantique (e.g., Adlam et al., 2009; Graham et al., 2000; Irish et al., 2016; Simons et al., 2001). La DS est la variante temporelle des démences fronto-temporales et elle se caractérise par une perte progressive des connaissances sémantiques résultant d'une neurodégénérescence des lobes temporaux (Hodges et al., 1992; Snowden et al., 1989; Warrington, 1975). Les patients atteints de DS présentent généralement des déficits

sur les tâches impliquant la MS comme les tâches de fluence catégorielle, de dénomination, d'association sémantique, et une préservation relative des autres fonctions cognitives comme la phonologie, la syntaxe, la résolution de problèmes non verbaux, les capacités visuo-spatiales et la mémoire au jour le jour (Crutch & Warrington, 2002; Irish et al., 2011; voir Hodges & Patterson, 2007). Si ces patients présentent des difficultés sur des tâches verbales de ME (Irish et al., 2016), des études ont montré qu'ils arrivaient tout de même à apprendre de nouveaux épisodes malgré leurs déficits sémantiques, en particulier, lorsqu'il était possible de se baser sur d'autres informations (e.g., perceptives). Par exemple, lorsque leur mémoire de visages inconnus qu'ils venaient de voir était évaluée par un test de reconnaissance, les personnes DS présentant une atrophie prédominante à gauche obtenaient des résultats similaires aux contrôles, bien que les patients avec atrophie bilatérale ou prédominante à droite avaient de moindres performances que les contrôles (Simons et al., 2001). Ces mêmes auteurs (Simons et al., 2001) ont également présenté deux cas de personnes DS avec atrophie prédominante à gauche à qui des visages de personnes célèbres étaient présentés. Les visages étaient divisés en deux catégories : ceux qui étaient identifiés par les personnes et ceux qui ne l'étaient pas. Lors de la phase de reconnaissance, les visages des mêmes personnes étaient représentés parmi des distracteurs mais soit la même image était utilisée (i.e., caractéristiques perceptives identiques), soit une image de la même personne mais dans une inclinaison différente (i.e., caractéristiques perceptives différentes). Les patients devaient sélectionner les visages des personnes qu'ils avaient vus dans la phase d'apprentissage, que l'image soit identique ou non. Les résultats ont montré que ces deux personnes DS avaient une bonne mémoire des visages, qu'ils aient été identifiés ou non, dans le cas où les mêmes images étaient utilisées à l'encodage et au test. Cependant, si des images différentes étaient utilisées, seuls les visages des personnes identifiées étaient bien reconnus mais non ceux des personnes non identifiées. D'autres études ont montré des performances préservées sur des tâches de ME chez ces personnes (Adlam et al., 2009; Irish et al., 2016). Cela suggère donc que même si les performances ne semblent pas optimales, il est tout de même possible d'avoir des compétences épisodiques en l'absence de traitement/ représentation sémantique, mais plus particulièrement si les informations perceptives peuvent être utilisées pour reconnaître les bonnes cibles.

Si les études conduites auprès de personnes présentant une altération de la MS semblent montrer que la MS n'est pas totalement nécessaire pour encoder ou récupérer des informations en ME, chez des individus sains, le traitement sémantique des informations joue un rôle dans la ME et peut aider à optimiser les performances en mémoire. Notamment, l'influence des traitements sémantiques sur la ME peut être mise en avant par plusieurs phénomènes. Tout d'abord, cela peut s'illustrer par la tendance à regrouper par catégorie

sémantique les éléments d'une liste de mots à rappeler (Bousfield et al., 1958). Par exemple, pour une liste comprenant des noms d'animaux, de légumes et de professions mélangés, une stratégie peut être de d'abord rappeler les éléments d'une première catégorie (e.g., animaux) puis ceux d'une seconde catégorie (e.g., légumes) et ainsi de suite. L'avantage du recours aux connaissances sémantiques sur les capacités de rappel peut également être montré par les travaux sur le rôle de l'expertise (Alba & Hasher, 1983; voir Gobet, 1998 pour une revue sur les joueurs d'échecs). Par exemple, un effet particulièrement documenté dans la littérature est que des joueurs d'échecs experts arrivent mieux à rappeler les positions des pions sur le plateau, lorsque ceux-ci sont cachés, que des joueurs novices. Cet effet serait en grande partie dû à la plus grande base de connaissances des coups à jouer chez les joueurs experts.

D'un autre côté, cette influence de la MS sur la ME peut aussi être démontrée à travers la possibilité de créer un faux souvenir en se basant sur le lien sémantique entre des items. Le paradigme Deese–Roediger–McDermott (Deese, 1959; Roediger & McDermott, 1995) montre qu'en demandant à des participants d'apprendre une liste de mots sémantiquement associés (e.g., lit, repos, éveil) à un mot distracteur non-présenté (e.g., sommeil), le distracteur était par la suite rappelé jusque dans 45% des cas et reconnu avec un haut niveau de confiance. Plusieurs hypothèses ont été proposées pour expliquer ce phénomène. L'hypothèse « associative » (Deese, 1959; Roediger et al., 2001) suggère que présenter ou rappeler plusieurs mots ayant un fort degré d'association avec le distracteur (i.e., mot non présenté mais souvent rappelé) activerait ce distracteur via un mécanisme d'amorçage automatique. D'un autre côté, d'après l'hypothèse « sémantique », ces distracteurs seraient activés car ils partagent des caractéristiques communes avec les mots présentés et ces caractéristiques seraient activés lors du traitement ou du rappel des mots (voir Brainerd et al., 2008; Gallo, 2006).

D'autre part, le degré de rétention d'une information dépend du niveau de traitement dont elle a fait l'objet. En effet, encoder des informations en faisant appel à leur signification permet un traitement profond et comporte un avantage sur la ME et la récupération future de ces informations par rapport à un encodage reposant sur un traitement superficiel (e.g., perceptif), même lorsque cette profondeur de traitement de l'information est inconsciente ( Craik & Lockhart, 1972; Lockhart & Craik, 1990). Après avoir traité un stimulus au niveau perceptif (e.g., lire un mot), approfondir ce traitement au niveau sémantique permet d'élaborer l'encodage et donc de renforcer la trace mnésique de ce stimulus. Par exemple, le mot « courgette » sera plus facilement retenu s'il est encodé par son sens (« est-ce que ce mot est un légume ? ») que si seules ses caractéristiques phonologiques sont activées (« est-ce que ce mot rime avec balayette ? »). Des théories ont été développées pour expliquer cet effet de profondeur d'encodage (e.g., théorie du double codage, D'Agostino et al., 1977, Anderson & Peterson, 2022).

Ensuite, des travaux conduits sur la mémoire associative ont montré que des mots présentés par paires étaient mieux retenus si les mots étaient sémantiquement reliés (e.g., fourmi-araignée) par rapport à des paires non reliées (e.g., fourmi-coussin) (Epstein et al., 1975; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003). Le lien sémantique permettrait de créer une association plus rapide et plus évidente entre les deux mots car elle se baserait sur des liens préexistants entre les concepts, et pourrait entraîner la possibilité d'un traitement plus profond. À l'inverse, des ressources supplémentaires seraient nécessaires pour créer une association en mémoire entre deux mots non sémantiquement reliés. Cet effet pourrait être expliqué par le recours à l'imagerie mentale (Bower, 1970), des processus d'intégration (Bein et al., 2015), et la profondeur d'encodage (Epstein et al., 1975).

Enfin, l'influence des connaissances préexistantes sur l'apprentissage de nouvelles informations épisodiques a également été étudiée par l'effet de congruence sémantique. La congruence sémantique peut être définie comme la congruence entre ce qui est perçu et ce qui est connu. Son influence sur la mémoire peut par exemple être évaluée en manipulant des associations de visages et de mots (e.g., Krebs et al., 2015), des présentations simultanées dans plusieurs modalités (e.g., Heikkilä et al., 2015), et des scènes visuelles (e.g., Ortiz-Tudela et al., 2017). La question de savoir si les informations qui sont congruentes avec les connaissances préalables sont mieux ou moins bien mémorisées que celles qui violent ces schémas préexistants est un débat d'actualité. Plus particulièrement, quelques études ont manipulé la congruence sémantique par l'utilisation d'images d'objets insérés dans des scènes visuelles sémantiquement congruentes ou incongruentes, et montré comment cela influençait les performances en mémoire à long-terme (Ortiz-Tudela et al., 2017). Par exemple, dans deux études (Expériences 2 et 3 de Ortiz-Tudela et al., 2017), les participants voyaient des images de contextes (i.e., scènes visuelles) seules puis dans certains essais, un item (e.g., vache) apparaissait sur l'image. Les contextes étaient congruents (e.g., champs) ou incongruents (e.g., route) avec les items. Les participants devaient indiquer le plus rapidement possible dès qu'ils détectaient que quelque chose avait changé sur l'image, puis essayaient d'identifier l'item (i.e., dire de quel item il s'agissait). Dans les cas où ils n'identifiaient pas les items, ils devaient indiquer l'endroit de l'image dans lequel l'item était apparu. Les participants étaient ensuite testés sur leur mémoire des items puis des scènes : plusieurs items (ou scènes) étaient présentés et pour chaque, ils devaient indiquer s'ils avaient été vus ou non dans la phase de détection. Les items vus dans des scènes congruentes étaient mieux mémorisés que les items vus dans des scènes incongruentes, mais aucun effet de congruence n'était observé sur la mémoire des scènes.

En conclusion, bien que le traitement sémantique des informations ne semble pas nécessaire à l'encodage et à la récupération d'informations en ME, les informations sémantiques, qui peuvent être activées automatiquement et pas nécessairement de manière consciente (Neely, 1991; Neely & Kahan, 2001), influencent tout de même la rétention des souvenirs épisodiques de multiples manières.

### 3.2.3. Utilisation de la mémoire sémantique dans la métacognition

Nous avons jusqu'ici montré que le traitement sémantique peut influencer l'encodage et la récupération en ME. Ainsi, il est aussi important de se demander si les personnes ont connaissance de ces effets et également, si elles arrivent à monitorer leurs performances sur des tâches dans lesquelles les informations sémantiques sont manipulées. De plus, on peut se questionner sur l'utilisation des informations sémantiques concernant les stimuli à mémoriser sur une tâche de ME pour évaluer les performances mnésiques.

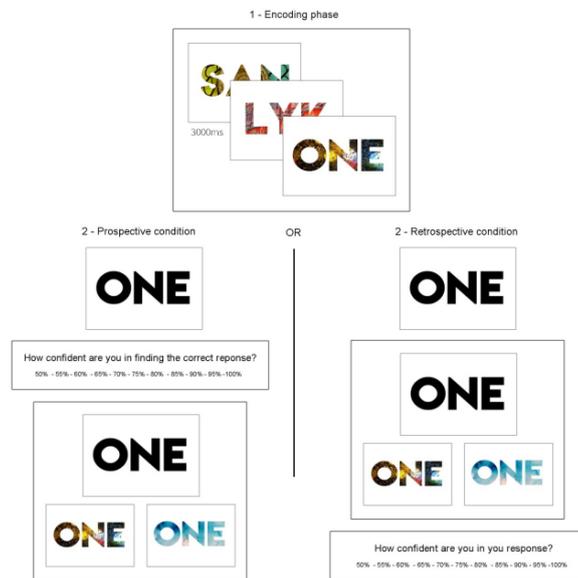
L'utilisation des connaissances sémantiques pour faire des jugements métacognitifs sur des tâches de ME peut être illustrée de plusieurs manières. Tout d'abord, comme introduit dans le [Chapitre 2](#), certains travaux ont étudié l'intervention de la fluence conceptuelle (i.e., facilité de traitement d'une information selon qu'elle soit pertinente avec nos connaissances préalables ou non) sur la formation des jugements métacognitifs épisodiques. Le principe de manipulation de la fluence conceptuelle est donc de jouer avec des facteurs qui reposent sur les représentations sémantiques (ou lien entre ces représentations) pour rendre l'information plus ou moins facile à traiter. Par exemple, des travaux ont manipulé la fluence conceptuelle par le degré d'association sémantique entre deux mots de paires à retenir. Cette manipulation était utilisée comme indice et influençait les prédictions de rappel faites juste après l'encodage de ces paires (i.e., JOL ; voir Mueller et al., 2013 pour une revue). Notamment, juste après avoir appris des paires des mots, les personnes attribuaient des JOL plus élevés aux paires sémantiquement reliées (e.g., pré – cheval) prédisant ainsi qu'ils pourraient mieux les rappeler par la suite par rapport aux paires non reliées (e.g., herbe - avion) (Castel et al., 2007; Connor et al., 1997; Dunlosky & Matvey, 2001; Undorf & Erdfelder, 2015). Cet effet était également observé lorsque les participants rappelaient en réalité plus de paires non reliées que de paires reliées (Carroll et al., 1997). Dans leur étude (Expérience 1), Carroll et al. (1997) ont contrasté le degré d'association sémantique des paires à leur degré d'apprentissage. Les participants apprenaient des paires non reliées et reliées puis devaient rappeler les cibles lorsque l'expérimentateur leur donnait l'indice. Lorsque les participants ne se rappelaient plus de la cible, la paire leur était remontrée. Les paires étaient réappries jusqu'à ce que les cibles soient rappelées 2 fois (pour les paires reliées) ou 8 fois (pour les paires non-reliées). Une fois

le critère d'apprentissage atteint pour toutes les paires, chaque indice était représenté et les participants prédisaient à quel point ils pensaient pouvoir rappeler la cible dans la phase de test (JOL différé). Deux ou 6 semaines après, chaque indice était représenté et les participants devaient rappeler les cibles associées. Pour les paires non rappelées, les participants devaient faire un FOK, puis choisir parmi 4 alternatives la bonne cible. Les JOL étaient plus influencés par le degré d'association (i.e., indice intrinsèque) alors que les FOK étaient influencés par le degré d'apprentissage (i.e., indice extrinsèque ou mnésique). En effet, les participants donnaient des JOL plus élevés pour les paires reliées alors qu'elles étaient en réalité moins bien rappelées que les paires non-reliées. À l'inverse, les FOK étaient plus élevés pour les paires non-reliées (sur-apprises) que pour les paires reliées (sous-apprises), ce qui reflétait le pattern de performances à la tâche de reconnaissance. Notons cependant que dans une seconde expérience (Expérience 2, Carroll et al., 1997), l'effet du degré d'association disparaissait lorsqu'un délai d'un jour était introduit entre la phase d'apprentissage et le moment où les JOL étaient faits. Ces résultats suggèrent que la fluence conceptuelle manipulée par le degré d'association joue un rôle plus important dans la magnitude des JOL que dans celle des FOK, et plus particulièrement lorsque qu'ils sont faits juste après que les paires aient été apprises. Une autre étude a montré que la fluence conceptuelle influençait également la sensibilité des JOL puisque ces jugements étaient moins corrélés aux performances réelles lorsque les mots étaient reliés que lorsqu'ils ne l'étaient pas (Hertzog et al., 2002).

En lien avec les travaux cités précédemment, plusieurs travaux ont documenté l'effet d'attente (i.e., *expectancy illusion*) sur le monitoring de source (i.e., mémoire de l'endroit où/personne par laquelle une information a été apprise). Par exemple, dans des expériences où des items (e.g., grille-pain) étaient associés à des mots de contexte attendus (e.g., cuisine) ou inattendus (e.g., salle de bain), les participants prédisaient généralement pouvoir plus se souvenir du contexte pour les items présentés avec un contexte attendu plutôt qu'inattendu (Schaper et al., 2019a, 2019b; Schaper & Bayen, 2021). Cependant, les performances en rappel reflétaient un pattern inverse : la mémoire de source était généralement meilleure pour les items ayant été appris dans des contextes qui violent les attentes (e.g., un grille-pain dans la salle de bain) que les items ayant été appris dans des contextes attendus (e.g., un grille-pain dans la cuisine). Cet effet est appelé effet d'inconsistance (voir Kuhlmann & Bayen, 2015 pour une revue).

De plus, une étude menée au sein de notre laboratoire et utilisant un autre paradigme a montré l'influence de la fluence conceptuelle sur les jugements métacognitifs épisodiques (Mazancieux, 2020). Les participants apprenaient des patterns visuels insérés dans des mots ou des pseudo-mots (voir Figure 3.5). Dans une seconde phase, les mots/pseudo-mots étaient représentés sans pattern, et pour chaque, un groupe de participants devait prédire s'ils

pourraient reconnaître le pattern qui allait dedans (i.e., jugements prédictifs) puis sélectionner parmi 2 propositions le bon pattern. L'autre groupe de participants devait d'abord sélectionner le bon pattern puis indiquer à quel point ils étaient certains d'avoir sélectionné le bon (i.e., jugement rétrospectif). La fluence conceptuelle était ici manipulée par la lexicalité des items présentés. L'efficacité métacognitive était meilleure pour les essais avec des mots (i.e., items fluents) par rapport aux essais avec des pseudo-mots (i.e., items non-fluents), mais seulement lorsque les jugements étaient rétrospectifs. Il n'y avait pas d'effet significatif de la lexicalité sur l'efficacité métacognitive pour les jugements prédictifs. Ainsi, cette étude montre que la sensibilité des jugements rétrospectifs étaient plus influencés par la fluence conceptuelle que celle des jugements prédictifs.



**Figure 3.5 :** Paradigme utilisé dans la thèse de Mazancieux (2020). La fluence conceptuelle était manipulée par la lexicalité des items dans lesquels les patterns étaient insérés.

L'influence du traitement sémantique sur la métacognition épisodique peut également être illustrée par d'autres études qui ont montré que les niveaux d'encodage jouaient un rôle dans les prédictions métacognitives (Expérience 1 de Lupker et al., 1991; Expérience 1 de Thomas et al., 2012 mais voir Expérience 2 pour des résultats contradictoires avec des images, dans lesquelles les informations perceptives). Dans une de ces études (Expérience 1 de Thomas et al., 2012), les participants reconnaissaient et prédisaient plus de reconnaissance (FOK) des cibles appartenant aux paires encodées en lien avec leurs caractéristiques sémantiques plutôt que des cibles des paires encodées au niveau perceptif. Les participants avaient également une meilleure sensibilité métacognitive (mesurée par une corrélation gamma) pour les paires encodées en profondeur. Dans une autre étude (Expérience 1 de Lupker et al., 1991), la

profondeur d'encodage était également manipulée au moment de l'encodage de paires : les paires étaient apprises en générant une phrase qui a du sens (i.e., condition d'encodage profond) ou en comptant le nombre de voyelles des mots (i.e., condition d'encodage superficiel). Les participants devaient ensuite essayer de rappeler les cibles sur présentation des indices. Lorsqu'ils n'arrivaient pas à les rappeler, ils devaient indiquer à quel point ils pensaient connaître l'item. Ici, les jugements métacognitifs étaient donc plus portés sur le présent puisqu'il n'y avait pas de demande explicite de prédire une performance future. Le test final comportait l'indice de chaque paire et les 3 premières lettres des cibles pour chaque. Les participants devaient compléter la suite de la cible. Même si les jugements métacognitifs ne portaient pas explicitement sur la performance future, ces jugements étaient plus corrélés à la performance à la tâche de complétion pour les paires apprises dans des conditions d'encodage profond par rapport aux paires apprises dans des conditions d'encodage plus superficiel. Ces deux études suggèrent donc que les personnes arrivent à bien monitorer l'effet de profondeur sur la performance en ME, mais également, que la valeur prédictive des FOK est meilleure lorsque les caractéristiques sémantiques des items sont mises en avant au moment de l'encodage en mémoire. Les traitements sémantiques semblent donc également jouer un rôle dans les jugements métacognitifs comme les prédictions faites au moment de la tentative de rappel.

Enfin, d'autres études menées sur le FOK ont montré que les caractéristiques sémantiques des items mémorisés servaient d'indice pour prédire la future récupération en mémoire (Thomas et al., 2012). Plus particulièrement, lorsque les cibles ne peuvent pas être rappelées, des informations partielles sur ces cibles sont parfois récupérées et influencent les prédictions de reconnaissance (i.e., FOK). Ces informations partielles sont parfois des attributs sémantiques (e.g., « c'était un animal ») (Koriat et al., 2003). Par exemple, Thomas et al. (2012; Expériences 1 et 2) ont montré que lorsque des informations sémantiques correctes étaient récupérées sur la cible juste après la tentative de rappel (e.g., dire si la cible était de la même catégorie sémantique que l'indice), les FOK étaient plus élevés pour les essais par la suite réussis que pour les essais ensuite échoués, ce qui suggère une certaine sensibilité métacognitive. Les FOK étaient également plus élevés quand les informations sémantiques récupérées étaient correctes que lorsqu'elles ne l'étaient pas, ce qui suggère que les personnes donnent plus d'importance à ces informations sémantiques lorsqu'elles sont justes. De plus, les FOK étaient plus élevés lorsque des informations sémantiques étaient récupérées que lorsque des informations perceptives (e.g., police d'écriture) étaient récupérées. Cela laisse suggérer que, pour guider leurs prédictions de reconnaissance, les personnes donnent plus de poids aux caractéristiques récupérées sur la cible lorsqu'elles sont de nature sémantique plutôt que perceptive.

En conclusion, toutes ces études montrent que les informations sémantiques sont utilisées pour guider les jugements métacognitifs. Notamment, l'influence du traitement sémantique sur ces jugements a été étudiée à travers des paradigmes manipulant la fluence conceptuelle (dans les JOL principalement), la profondeur d'encodage (dans le FOK), ou encore la récupération d'informations partielles sémantiques (dans le FOK).

### 3.3. Conclusion du Chapitre 3

Dans ce chapitre, nous avons vu que lorsque l'on s'intéresse aux compétences métacognitives, un facteur à prendre en compte est les capacités mnésiques mises en jeu par la tâche de premier ordre mais aussi lors de l'émission des jugements métacognitifs. En effet, cette relation entre tâche de premier ordre et celle de second ordre dépend des mesures utilisées pour évaluer la métacognition mais aussi des processus mnésiques mis en jeu sur la tâche de second ordre. De plus, nous avons vu que même s'il est pertinent de s'intéresser à la distinction épisodique/sémantique sur les tâches de métacognition, il est aussi important de considérer l'influence des traitements sémantiques sur la mémoire épisodique, mais également leur utilisation dans les tâches de métacognition portant sur la mémoire épisodique. En particulier, nous avons montré que les traitements sémantiques des informations étaient utilisés pour plusieurs types de jugements métacognitifs (e.g., fluence conceptuelle, profondeur d'encodage, informations partielles). Cependant, ces traitements sont-ils nécessaires pour pouvoir permettre d'évaluer les performances en mémoire ? De plus, la majorité des études portant sur la fluence conceptuelle ont porté sur le paradigme de JOL. Qu'en est-il de l'influence de cette fluence conceptuelle sur les jugements prédictifs faits au moment du test et sur les jugements rétrospectifs ? De plus, l'effet de la fluence conceptuelle sur la métacognition a majoritairement été étudié avec du matériel verbal (e.g., mots). Quelques études ont utilisé du matériel plus visuel (e.g., paires d'images dans Expérience 2 de Thomas et al., 2012) mais à notre connaissance aucune étude utilisant des scènes naturelles n'a été publiée.

### CHAPITRE 3 : POINTS IMPORTANTS

- Il existe une **relation** entre performance mnésique et performance métacognitive au niveau de la **mesure** et au niveau **cognitif**.
- Un déficit mnésique ne s'accompagne pas systématiquement d'un déficit métacognitif, même pour les types de jugements qui sont généralement les plus corrélés aux capacités mnésiques (e.g., FOK).
- Les processus de mémoire interviennent dans les jugements métacognitifs, et en particulier dans les prédictions comme les JOL et les FOK. Notamment, le traitement sémantique des informations à retenir en ME guide ces jugements dans certains cas.
  - Dans cette thèse, l'influence d'indices sémantiques sur les compétences métacognitives a été évaluée (Chapitres 6 et 7).

# Problématique générale

---

Dans les chapitres précédents, nous avons montré que plusieurs types de jugements métacognitifs pouvaient être utilisés pour évaluer les capacités de monitoring (e.g., prédictifs, rétrospectifs, épisodiques, sémantiques). Ces jugements mettent en jeu différents processus et se basent sur des indices divers. Ainsi, ce qui est observé dans un paradigme avec un type de jugement ne s'applique pas forcément aux autres paradigmes, que ce soit par rapport au moment où les jugements sont faits (i.e., prédictifs ou rétrospectifs) mais également au type de tâche de premier ordre (i.e., mémoire épisodique ou mémoire sémantique). De plus, nous avons également montré que la magnitude des jugements métacognitifs, mais également leur sensibilité, étaient dans certains cas influencées par des processus mnésiques épisodiques (e.g., recollection) ou sémantiques (e.g., fluence conceptuelle, profondeur d'encodage). Le but général des projets expérimentaux de cette thèse était d'approfondir la compréhension des processus mnésiques intervenant dans les jugements métacognitifs, en étudiant la relation entre performances de premier ordre et de second ordre, mais aussi en étudiant le rôle des traitements sémantiques dans les jugements métacognitifs.

Tout d'abord, pour certains jugements comme le FOK, il semble exister, dans plusieurs populations, un fractionnement de la métacognition. C'est notamment le cas dans le vieillissement, puisque certaines études suggèrent un déclin de la sensibilité des eFOK malgré une préservation du sFOK. Cependant, toutes les études ne convergent pas vers cette conclusion. Un premier objectif de cette thèse était de résumer les études publiées sur cette question et de mieux comprendre les disparités observées ([Chapitre 4](#)).

Ensuite, un grand nombre d'études a montré un déclin de la sensibilité des eFOK chez des personnes présentant un déclin de la mémoire épisodique mais peu ont rapporté les performances métacognitives dans le cas de déficit sémantique général. De plus, un déficit de premier ordre n'entraîne pas systématiquement un déficit métacognitif, de second ordre. Ainsi, un second objectif était d'évaluer les performances de personnes présentant une démence sémantique, une pathologie dégénérative touchant principalement la mémoire sémantique, sur plusieurs tâches de métacognition (i.e., jugements locaux prédictifs et rétrospectifs, sur des tâches épisodiques et sémantiques ; [Chapitre 5](#)).

Dans le [Chapitre 6](#), nous avons avancé l'idée que le système sémantique fournit non seulement un cadre à la ME mais également aux jugements métacognitifs prédictifs faits sur une tâche de ME. Ainsi, dans ce chapitre, nous avons étudié l'aptitude des personnes à prédire leur capacité future à reconnaître des informations apprises qui n'ont pas de sens (pseudo-mots). Nous cherchions à évaluer si avoir une représentation sémantique préexistante des informations à mémoriser est nécessaire pour pouvoir prédire de façon correcte leur reconnaissance future.

Enfin, plusieurs études ont montré que la fluence conceptuelle pouvait influencer les jugements métacognitifs. Cependant, la majorité de ces études ont porté sur les JOL et utilisé du matériel verbal pour manipuler la fluence. Ainsi, dans un dernier chapitre expérimental ([Chapitre 7](#)), nous avons approfondi la compréhension de l'influence des connaissances sémantiques sur les jugements métacognitifs prédictifs faits au moment du test et rétrospectifs en manipulant la fluence conceptuelle à l'aide d'images de scènes visuelles plus perceptivement riches que des mots.

L'hypothèse générale de cette thèse était que des facteurs mnésiques spécifiques interviennent et influencent les jugements métacognitifs (i.e., processus épisodiques influencent jugements métacognitifs sur une tâche de mémoire épisodique). Nous supposons donc une mémoire-spécificité des jugements métacognitifs comme les prédictions de reconnaissance. De plus, compte tenu de l'importance de la mémoire sémantique dans la cognition et de l'interaction entre mémoire sémantique et mémoire épisodique, nous faisons également l'hypothèse que des facteurs sémantiques peuvent influencer les jugements métacognitifs, y compris sur des tâches épisodiques. Nous faisons plusieurs hypothèses plus spécifiques. Tout d'abord, nous prédisions que la dissociation épisodique/sémantique qui s'observe sur certaines tâches métacognitives dans certaines populations découle, du moins en partie, de facteurs mnésiques et plus particulièrement du déclin en mémoire épisodique ([Chapitre 4](#)). Deuxièmement, nous faisons l'hypothèse qu'un déficit en mémoire sémantique pourrait également influencer la sensibilité métacognitive sur des tâches portant sur la mémoire sémantique, plus que sur celles portant sur la mémoire épisodique ([Chapitre 5](#)). Ensuite, puisque plusieurs études ont montré l'influence des traitements sémantiques sur la métacognition épisodique, nous présumons que les capacités métacognitives seraient réduites dans le cas où les personnes ne possèdent pas de représentations sémantiques préexistantes des informations à retenir ([Chapitre 6](#)). Enfin, nous faisons l'hypothèse que les traitements sémantiques peuvent également jouer sur des prédictions de reconnaissance et des jugements rétrospectifs épisodiques, à travers la manipulation de la fluence conceptuelle. Ces connaissances seraient utilisées comme indice pour les jugements métacognitifs ([Chapitre 7](#)).

# **PARTIE EXPÉRIMENTALE**

Les 4 chapitres expérimentaux ont tenté de répondre à différentes problématiques (voir Tableau I). Les Chapitres 4 et 5 présentent des études portant sur des données de personnes présentant un déclin de la mémoire : des personnes âgées saines ([Chapitre 4](#)) et deux populations cliniques (la démence sémantique et la maladie d'Alzheimer ; [Chapitre 5](#)). Les [Chapitre 6](#) et [7](#) portent sur le fonctionnement normal de sujets jeunes. Les méthodes des études des Chapitres 4, 6 et 7 ont été pré-enregistrées sur l'Open Science Framework (voir la section Méthode de chacun des chapitres pour le lien).

**Tableau I** : Récapitulatif des problématiques et des types de jugements utilisés dans chaque chapitre.

Chapitre	Problématiques	Hypothèses	Expérience	Jugement
Chapitre 4	<p><b>Lien entre type de déficit de premier ordre et capacités métacognitives</b></p> <p><i>Est-ce que les capacités mnésiques expliquent le pattern de détérioration/préservation de sensibilité métacognitive du eFOK et sFOK chez les personnes âgées ?</i></p>	Le déclin de la sensibilité du eFOK dans le vieillissement découle du déclin en mémoire épisodique	2 méta-analyses	eFOK/ sFOK
Chapitre 5	<p><b>Lien entre type de déficit de premier ordre et capacités métacognitives</b></p> <p><i>Quelles sont les capacités métacognitives des personnes présentant une pathologie neurodégénérative altérant majoritairement la mémoire sémantique ?</i></p> <p><i>Existe-t-il une relation entre type de déficit principal (i.e., épisodique ou sémantique) et sensibilité métacognitive ?</i></p>	Les personnes DS auraient une moins bonne sensibilité sur tâches sémantiques, surtout pour les PoR	1 expérience	ePoR/ sPoR/ sRCJ/ eRCJ
Chapitre 6	<p><b>Rôle du traitement sémantique dans la métacognition épisodique</b></p> <p><i>Est-ce que le traitement sémantique des informations est nécessaire pour prédire ses performances ?</i></p>	Traitement sémantique fournit un cadre aux PoR : moins bonne sensibilité pour prédire la récupération d'informations qui n'ont pas de sens	4 expériences : paires de mots et pseudo-mots	ePoR
Chapitre 7	<p><b>Rôle du traitement sémantique dans la métacognition épisodique</b></p> <p><i>Est-ce que les connaissances antérieures influencent les jugements métacognitifs et perturbent la capacité à s'évaluer même si la mémoire sémantique n'est pas directement évaluée par la tâche ?</i></p>	Congruence sémantique utilisée comme indice : confiance plus faible lorsque la cible est traitée de manière moins fluente	Expérience 7.1 : manipulation de la fluence conceptuelle au test  Expérience 7.2 : manipulation de la fluence conceptuelle à l'encodage	eRCJ  ePoR

# Chapitre 4. Distinction entre feeling-of-knowing épisodique et sémantique dans le vieillissement

Comme nous avons commencé à l'introduire dans les chapitres théoriques, dans la littérature, un pattern complexe de préservation et de détérioration des capacités métacognitives dans le vieillissement sain est retrouvé, notamment en ce qui concerne les prédictions d'une récupération future en mémoire (i.e., *feeling-of-knowing*, FOK). Alors que le FOK sémantique (sFOK) est préservé avec l'âge, les études sur les tâches épisodiques (eFOK) fournissent des résultats ambigus. Dans ce Chapitre, nous présentons une méta-analyse de 20 études sur l'eFOK et le sFOK, analysant la différence de sensibilité métacognitive entre 922 adultes jeunes et 966 adultes plus âgés, en prenant en compte la différence de performance en mémoire. Les études sur le sFOK n'ont révélé aucun effet de l'âge (8 effets,  $g = -0,10 [-0,29, 0,10]$ ). Cependant, nous avons trouvé une différence entre les groupes d'âge pour l'eFOK (22 effets,  $g = 0,53 [0,28, 0,78]$ ), qui était modérée en prenant en compte la performance en reconnaissance. De plus, en utilisant les données agrégées de 134 jeunes adultes et 235 adultes plus âgés provenant d'études publiées et non publiées de notre laboratoire, nous avons évalué le rôle de la performance en mémoire dans le déficit du eFOK. Nous montrons que les personnes âgées ont une moins bonne sensibilité métacognitive que les plus jeunes sur le eFOK, et que cela est lié au déclin de la mémoire avec l'avancée en âge. Nous mettons en évidence deux possibles explications non exclusives : un déclin de la mémoire en jeu dans les tâches de premier et de second ordre, et une contamination de la performance de premier ordre dans la mesure utilisée pour évaluer la sensibilité métacognitive.

Ce chapitre fait l'objet d'un article soumis à la revue *Scientific Reports*. Il a été conservé dans sa langue original d'écriture (anglais) et reformaté pour les besoins de ce manuscrit.

**Devaluez, M.\***, Mazancieux, A.\*, & Souchay, C. (2022). Episodic and semantic feeling-of-knowing in aging: a systematic review and meta-analysis.

\*co-auteurs

Quelques parties de ce chapitre ont été présentées : au Workshop "Metacognition and Aging" de l'Université de Padova, Italie, Avril 2021 ; au meeting de l'Experimental Psychology Society, Londres, UK, Janvier 2022 ; à la Journée d'Étude du Vieillissement, Lyon, Mai 2021.

## 4.1. Introduction

Metacognition, the ability to monitor and control our cognitive ability, is multifaceted. It has been investigated in the context of several domains (e.g., episodic memory, semantic memory, perception) or on several levels (local or global, Seow et al., 2021). Deficient metacognition has been shown to occur in various populations from people with neurological diseases (Pannu & Kaszniak, 2005) to psychiatric populations (Hoven et al., 2019). In aging, the question whether metacognition is altered is controversial. Although the alteration of frontoparietal networks in aging is known to involve regions implicated in metacognition (Fleming & Dolan, 2012; Vaccaro & Fleming, 2018), empirical evidence for such a deficit remains inconsistent. The question of whether metacognition is intact in aging is of importance, as it would be the basis for older adults to use compensatory strategies for any cognitive decline.

We aim to address the status of metacognition in healthy aging by focusing on a meta-analysis for one specific task - the feeling-of-knowing (FOK). The FOK is the ability to predict future recognition of currently unrecalable information. Numerous studies have shown that older adults make inappropriate evaluations of their memory for recently learned information on such a task. It is of particular interest since in most other experimental paradigms older adults and younger adults show equivalent metacognition. Age invariance (or even superior performance in older adults) in metacognitive sensitivity (i.e., the ability to discriminate between correct and incorrect responses) has been reported for retrospective confidence tasks (Dodson et al., 2007; Marquié & Huet, 2000; Pliske & Mutter, 1996), and for judgement-of-learning (JOL) and delayed-JOL paradigms (Connor et al., 1997; Hertzog et al., 2002; Tauber & Rhodes, 2012a). Similarly, recent work has shown preserved confidence judgments in this population (McWilliams et al., 2022). Thus, older adults' performance on episodic FOK (eFOK) tasks is of interest both for how we understand the cognitive aging process, but also how we conceive metacognition as a general function beyond a specific type of metacognitive judgment.

This idea that metacognition refers to a general function relates to the notion of the domain-generality of metacognition. From this perspective, metacognition is not entirely encapsulated in each cognitive domain but some general abilities are shared across cognitive domains (Carpenter et al., 2019; Mazancieux, Fleming, et al., 2020). Within FOK judgements, a contrast is made between semantic FOK (sFOK) and eFOK. In the eFOK, there is a first study phase using cue-target pairs followed by a FOK phase where the participants recall target words when given cues and/or make a prediction of future recognition. In sFOK tasks, there is no study phase but only an immediate FOK phase where the participant attempts to recall answers to general knowledge questions. In both tasks the accuracy of the FOK predictions is assessed by comparison with final recognition memory.

The accuracy of older adults' metacognitive evaluations on episodic tasks can either be contrasted with younger adults' performance on the same episodic task, or by comparing episodic and semantic materials. The claim is often made that whereas there is age-equivalence on the sFOK task, there are age differences on eFOK tasks (Douchemane et al., 2007; Morson et al., 2015; Souchay et al., 2007). This is found across several different neuropsychological populations (e.g., in autism spectrum disorder, Grainger et al., 2014; Wojcik & Souchay, 2013; in Alzheimer's disease, Souchay, 2007). Similarly, in a large sample of university students it was shown that whereas the sensitivity of retrospective confidence judgements was correlated across semantic and episodic tasks, the same pattern broke down for FOKs - a challenge for the concept of domain generality of metacognition (Mazancieux, Fleming, et al., 2020). Comparing FOKs for semantic and episodic material, a neuroimaging study inspired by the neuropsychological approach has shown different patterns of neural activations for the two FOK judgements (Reggev et al., 2011).

A few critical theoretical issues concerning metacognition in older adults and the domain generality of metacognition therefore rest on the finding that eFOK is differentially impaired in older adults. However, this is far from an unequivocal finding, with some studies suggesting age equivalence in eFOK (Eakin et al., 2014; Eakin & Hertzog, 2012; Hertzog et al., 2010; Souchay et al., 2002) and some studies reporting age differences (Sacher et al., 2015; Souchay et al., 2000; Souchay & Isingrini, 2012). One challenge is to better understand the discrepancies in the literature. Overall, studies addressing the eFOK sensitivity in relation to aging used similar paradigms and populations of interest, but they also show differences in methods (e.g., type of material, study time, FOK scale) and diverse conclusions that we review in the results' section. Of great importance here, some authors have taken additional steps to ensure equivalent memory performance between age-groups while others did not try to control for first order performance.

This difference is particularly relevant as poorer episodic memory is one of the main hypotheses that is proposed to account for the eFOK deficit in aging (Hertzog et al., 2010). In fact, multiple theories have been proposed to explain the potential specific impairment of eFOK in aging. One theory postulates that the age-related eFOK sensitivity and episodic memory impairments are driven by an executive function deficit (West, 1996). Support for this theory comes from results showing a correlation between eFOK sensitivity and executive function test scores (Perrotin et al., 2006; Souchay et al., 2000) and regression analyses suggesting executive functioning as a main factor of eFOK accuracy (Perrotin et al., 2008). These results are in line with functional neuroimaging results highlighting the importance of a fronto-temporal network and the critical involvement of the prefrontal cortex in eFOK (Schnyer et al., 2005), which might underlie a monitoring process involving interactions between executive and memory functions.

A deficit in self-initiated processes has also been proposed as the basis for the decrease in eFOK sensitivity with aging (Thomas et al., 2012). These authors propose that the quality of partial information accessed about the target when making the FOK judgment is a determinant of FOK sensitivity. Older adults displayed better sensitivity when they had access to correct partial information. Thomas et al. (2012) suggest that the eFOK deficit observed in aging is thus associated with less effective use of partial information. Finally, two other theories advance that memory processes underpin the eFOK decline with aging. The memory constraint hypothesis (Hertzog et al., 2010) proposes that the decrease in FOK sensitivity on episodic memory tasks is a consequence of the quality of original encoding. Poorer encoding leads to weaker memory strength and does not allow enough information for accurate FOKs. Other work (Souchay et al., 2007) suggests that this eFOK deficit is a consequence of a lack of recollection (of contextual or retrieval cues) during the recall attempt made at the moment of the FOK. Older adults make less accurate predictions because they fail to recollect the target or any information that could lead them to believe they know the target. This result has been recently strengthened by a study showing that metacognition efficiency (i.e., metacognitive sensitivity that controls for recognition memory performance) is correlated with recall performance during the FOK phase (Mazancieux, Fleming, et al., 2020).

As memory processes contribute to the eFOK deficit, a challenge is to examine whether the observed eFOK deficit in aging is a direct consequence of the episodic memory decline associated with age. The relationship between first order (i.e., memory) and second order (i.e., metacognitive) performance is a critical point and a source of debate in the field of metacognition. Several classical measures of metacognitive sensitivity have proven to correlate with first order performance (e.g., Goodman-Kruskall gamma correlation; see Fleming & Lau, 2014 for a review). As such, the question raised above becomes particularly interesting given that the literature about FOK in aging relies strongly on the gamma correlation (Goodman & Kruskal, 1954; Nelson, 1984). Gamma correlations measure the degree of relationship between the accuracy of a response given during the first order task and the metacognitive judgment for the same trial. Therefore, a strong unbalance between correct and incorrect responses in the first order task biases the gamma correlation. Similarly, a high proportion of high or low FOK responses also strongly modifies its value. This second issue suggests more specifically that the eFOK deficit observed in aging could also be the result of a metacognitive bias different from young adults' (i.e., over- or under-confidence).

We present a systematic review of published studies conducted on eFOK and sFOK in aging. In the first part, using a meta-analytic approach, one goal was to assess the general age-related performance on both episodic and semantic tasks in the existing literature. The issue of a confound between eFOK sensitivity and memory performance is not a new concern,

and as such several published studies have taken additional steps to limit memory inequality between age groups, by either constraining younger adults or favoring older adults' memory performance. We explored whether this specific deterioration in eFOK sensitivity might be a consequence of episodic memory impairment as it has been done for other populations (e.g., Rouy et al., 2021 in schizophrenia). In a second part, using an aggregated dataset from published and unpublished studies conducted in our lab, we aimed at providing additional evidence of the contribution of first order performance to metacognitive sensitivity by matching participants according to memory performance using either recall or recognition. Knowing the limits of standard measures of FOK sensitivity, we also intended to evaluate eFOK sensitivity in older and younger adults using classical gamma correlations as well as measures less contaminated by metacognitive bias (e.g., type-2 d'). In both parts, we expected a reduced difference in eFOK sensitivity between older and younger adults when first order performance (recall or recognition) was controlled for than when it was not considered.

## 4.2. Methods

The method and analyses were pre-registered on the Open Science Framework (<https://osf.io/5hkpt/>). Raw data extracted from each selected article and analysis scripts are available on GitHub ([https://github.com/amazancieux/eFOKaging\\_review](https://github.com/amazancieux/eFOKaging_review)).

### 4.2.1. Systematic review and global meta-analysis

#### *Selection and inclusion*

The systematic review and meta-analysis were conducted following the PRISMA guidelines and recommendations (PRISMA-P Group et al., 2015). Summary of the selection steps are described in

Figure 4.1.

**Identification.** Records published as of October 15th 2020 were identified from PubMed and Web of Sciences online databases. Two searches were carried out each using the keyword “feeling-of-knowing” associated first with “older adults” and then with “aging”. Additional reports were identified by checking references in selected papers. No time limit was set regarding the year of publication. After records were identified, duplicates were removed.

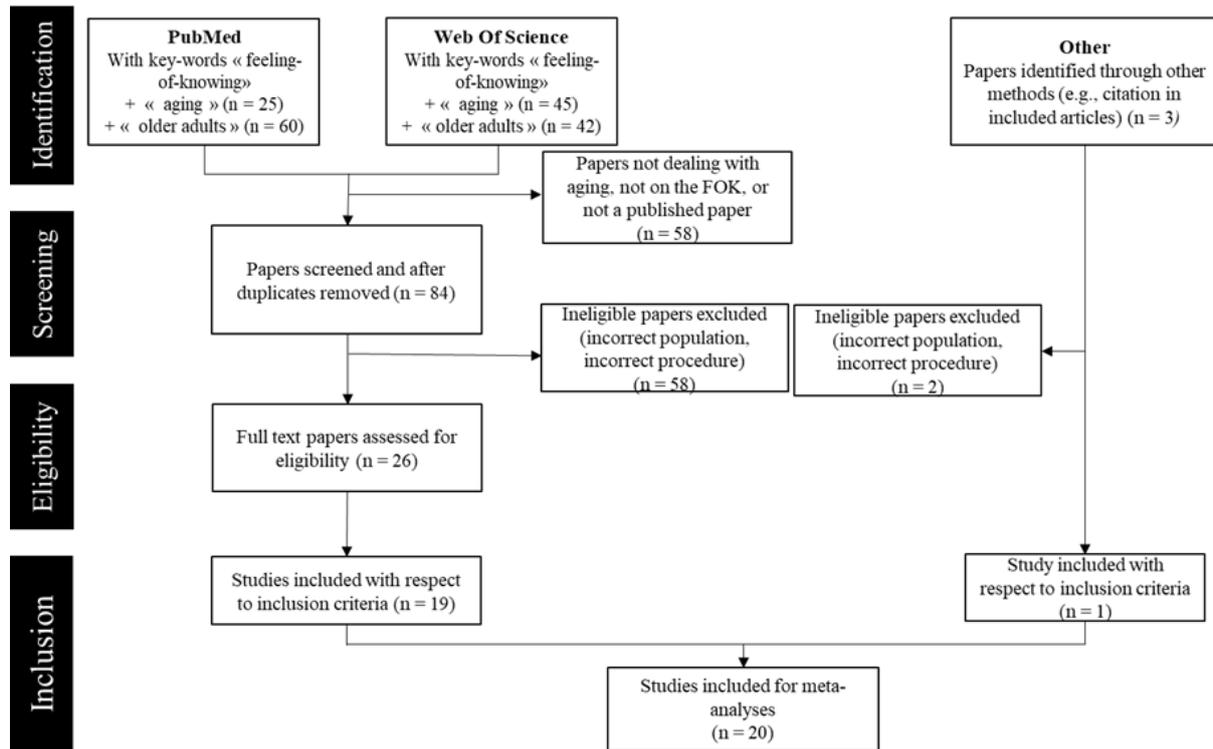
**Screening.** Articles were first selected on the basis of their title and abstract. Records not dealing with aging or not using the FOK procedure were excluded. We also withdrew records which were not published papers or not original research reports.

**Eligibility criteria.** For inclusion in the systematic review, full texts of selected articles were inspected. Only records which met eligibility criteria were included. To be eligible, records

had to be original research articles written in French or in English. Inclusion criteria also comprised comparison of a group of younger adults (YA) and a group of older adults (OA) using a sFOK or eFOK standard paradigm. In an eFOK paradigm, participants predict the likelihood of future recognition of newly learnt material (e.g., pictures or words). We excluded studies in which the metacognitive judgment was called a 'feeling of knowing' but did not consist in a prediction of future recognition performance (e.g., Bäckman & Karlsson, 1985; Experiments 2 and 3 of Allen-Burge & Storandt, 2000) or was not a standard FOK procedure (Hertzog & Touron, 2011). To be eligible, studies also had to describe performance using a measure of metacognitive sensitivity (e.g., gamma correlation, Hamman correlation, type-2 d'). In order to perform statistical analyses, an additional eligibility criterion for inclusion in the meta-analysis was the description of sample sizes, means and standard deviations for the metacognitive sensitivity measure.

As introduced briefly above and discussed in detail below, several studies included multiple comparison groups with the aim of equating first order performance between YA and YA. Sacher et al. (2013) includes three YA groups and one OA group. Because all three YA groups are of interest for our main memory hypothesis on eFOK, we decided to include them all in the meta-analysis. To consider the fact that these 3 comparisons include the same OA group, we used a hierarchical meta-analytic model that takes into account the 'effect' variability within a particular study (see [Statistical analyses section](#) below for more details). Hertzog, Dunlosky & Sinclair (2010) compared two groups of older adults tested either after a 48-hour or 30-minute delay with one group of younger adults tested after a seven-day delay. However, as data for the 30-minute condition was not available in the article, we decided to exclude this condition from the current meta-analysis.

Several studies also included other within-subject manipulations but we did not focus on these manipulations. In such studies (Thomas et al., 2011), we included the overall performance instead of data for each specific condition. Thomas et al. (2011) also added between-subject manipulations. In their Experiment 2, participants attempted recall of partial information either before or after the FOK. In Experiment 3, participants had a time limit to either perform their FOK or to retrieve partial information. For each experiment, data for both conditions were included separately in the analysis. In Eakin & Hertzog (2012), overall performance was not provided. Data for each condition was used and the comparison between the two groups was made for each condition. Finally, Eakin & Hertzog (2012) proposed 2 procedures (intralist and extralist cueing conditions). As this extralist condition moves away from the classical FOK paradigm, we decided to not include this data in the meta-analysis.



**Figure 4.1:** Flow chart of the steps for identification and selection of records included in the systematic review and meta-analysis.

### Statistical analyses

Analyses were performed on the measure of metacognitive sensitivity using four models (M1 to M4). All studies used gamma correlations as a measure of sensitivity. Hedges’s G was calculated to measure the effect size of age on metacognitive sensitivity. All analyses were performed using R software and multilevel meta-analysis models were carried out with the *metafor* package.

Our first meta-analytic model (M1) estimated the overall effect-size of a difference in metacognitive sensitivity between YA and OA taking also variability into account at three levels: the study level, the experiment level, and the effect level. We ran two versions of this model: one for eFOK (eM1) and one for sFOK (sM1).

As preregistered, we tested the hypothesis that eFOK effect can be reduced by controlling for memory. Two models were created: one testing recognition performance as a moderator and the other testing recall performance as a moderator, as they both capture similar but also different aspects of the confounds. Controlling for recognition performance allows an estimation of older adults’ eFOK deficit that could result from both (1) the intrinsic relationship between episodic memory and metamemory (i.e., metamemory sensitivity is based on memory processes such as recollection) and (2) confounds in statistical/mathematical quantification of

metacognitive sensitivity (i.e., the use of gamma as it directly takes into account recognition performance in its calculation). On the other hand, as FOKs are performed after a recall attempt, controlling for recall would control for the intrinsic relationship between episodic memory and metamemory. Indeed, FOKs are influenced by the ability to recall the target or specific information about it. As an example, Mazancieux, Dinze, et al. (2020) found a correlation between recall and metacognitive efficiency in both eFOK and sFOK. In short, whereas recognition performance may contribute to a statistical artifact in measures of metacognitive sensitivity as well as looking at genuine mnemonic factors, the recall measure captures something slightly different, since it does not contribute directly to the measure of metacognitive sensitivity, and captures memory function at the point of making the FOK judgment, not subsequent to it.

For the M2 model, we calculated Hedge's  $g$  effect sizes for the difference between YA and OA in recall (using means and standard deviations). This model was identical to M1e but included recall effect sizes as a moderator. Because of missing information in several articles, only 16 out of 22 effects (630 YA and 722 OA) were included in this analysis. For the M3 model, the same procedure was used with recognition performance including 14 out of 22 effects (536 YA and 606 OA). Finally, our last preregistered moderator to be tested was the type of recognition (e.g., 2AFC, yes/no). For the M4 model, we added to the M1e model a nominal moderator corresponding to the type of recognition task used for each effect.

All moderator analyses (M2, M3, M4) were performed in case of significant heterogeneity in M1e. This assessment was carried out using the  $Q$ -statistic (Card, 2015) and the  $I^2$  index, which corresponds to the percentage of the total variation due to between-studies variability (Higgins & Thompson, 2002).  $I^2$  values above 50% are considered a large amount of heterogeneity.

Publication bias was assessed using funnel plots (one for eFOK and one for sFOK) of observed outcomes according to corresponding standard errors (Sterne & Harbord, 2004). The plot asymmetry was tested using an adapted Egger's test for multilevel models.

#### **4.2.2. Analyses of individual scores from aggregated dataset**

##### ***Dataset description***

Data was taken from 5 published (Perrotin et al., 2006; Souchay et al., 2000; Souchay & Isingrini, 2012; two studies from Souchay et al., 2007) and 1 unpublished studies conducted in our lab (note that the 5 published studies were also included in the global meta-analysis). Data included recall performance and counts of recognized and unrecognized items given a yes or no FOK, producing a 2x2 table (see Tableau 4.1). This structure allowed us to calculate

gamma correlation (for our preregistered confirmatory hypotheses) but also Hamman coefficient and type-II  $d'$  (according to our preregistered exploratory hypotheses). Because several participants had no data in at least one of the boxes of 2x2 of the table, we calculated the corrected gamma, Hamman, and type-II  $d'$ . Note that we found some discrepancies between calculated means and means reported in corresponding articles as a probable consequence of disparities in methods used for correction of scores or rounding methods.

This aggregated dataset of individual data allows us to have more sensitive tests of our hypothesis. Therefore, we aimed to (1) reproduce findings of the global meta-analysis, (2) compare gamma across 50% of the OA with the best recall performance and 50% of the YA with the worst recall performance, (3) compare gamma for OA and YA with recall performance as a covariable, and (4) explore the effect of type of metacognitive sensitivity measure. For our confirmatory hypothesis (points 2 and 3), we expect to find a lower eFOK difference between YA and OA.

**Tableau 4.1** : Available data for each individual of the aggregated dataset.

Corrected gamma scores were calculated using  $\gamma = \frac{(A*D)-(B*C)}{(A*D)+(B*C)}$

	FOK Yes	FOK No
Correct recognition	A	B
Incorrect recognition	C	D

### **Statistical analyses**

Analyses were also performed using the R software with the *metafor* package. Hedges's  $G$  was calculated to measure the effect size of age on metacognitive sensitivity for each study. Several meta-analytic models were performed. The first (M5) estimated the overall effect-size of a difference in eFOK metacognitive sensitivity between YA and OA as a reproduction of the global meta-analysis. Model M6 estimated the overall effect-size of a difference in recall performance between YA and OA. Then, the following models were run in order to test the influence of memory performance in metacognitive sensitivity. Models M7a and M7b were performed on half of the sample size. For each study both YA and OA were split into 2 groups according to their recall performance (median split). M7a compared eFOK for OA with the highest recall performance and YA with the lowest recall performance. On the contrary M7b compared eFOK for OA with the lowest recall performance and YA with the highest recall performance. As preregistered, we supposed a reduction of the overall estimated effect in M7a compared to M7b.

## 4.3. Results

### 4.3.1. Systematic review and global meta-analysis

#### *Qualitative review*

In total, 22 effects taken from 20 experiments were included. [Appendix A](#) provides a summary of the studies investigating FOK sensitivity in healthy aging which were included. Thirteen studies explored the age-related effect on the eFOK only, 3 on the sFOK only, and 4 considered both the eFOK and sFOK. All studies compared at least one OA group to one YA group. Some studies included multiple groups of one age-category because authors tried to compare groups both when memory performance was equated and when it was not.

In the papers selected for review, 12 studies described a significant age-related deficit in FOK sensitivity (Douchemane et al., 2007; Morson et al., 2015; Perrotin et al., 2006; Sacher et al., 2013, 2015; Souchay et al., 2000; 2 experiments in Souchay et al., 2007 ; 3 experiments in Thomas et al., 2011). Two of these reports also showed reduced or no significant age-related difference in eFOK sensitivity under conditions allowing equated memory performance between YA and OA groups (Sacher et al., 2013) or when memory performance (recall or recognition) was controlled for in the analyses (Souchay & Isingrini, 2012). On the contrary, 4 studies reported a significant age effect on eFOK sensitivity despite memory performance being equated between age groups (Morson et al., 2015; Experiments 1 & 2 of Thomas et al., 2011 for equal performance on both recall and recognition; Experiment 3 of Thomas et al., 2011 for equal performance only on recognition).

Additionally, 4 studies reported no significant age effect on eFOK sensitivity (Eakin et al., 2014; Eakin & Hertzog, 2012; Hertzog et al., 2010; MacLaverly & Hertzog, 2009). Memory performance was equated between YA and OA in one of those studies (Hertzog et al., 2010). Hertzog et al. (2010) manipulated delay between encoding and test for both age-groups allowing less delay to OA (48 hours for one group and 30 minutes for the other) than to YA (7 days). The two groups which had matched episodic recognition performance (48-h delay OA and YA group) did not show significant difference in eFOK sensitivity. The number of times items were presented at encoding was also manipulated. Notably, for items only presented once, YA showed better eFOK sensitivity than OA, despite memory performance being matched.

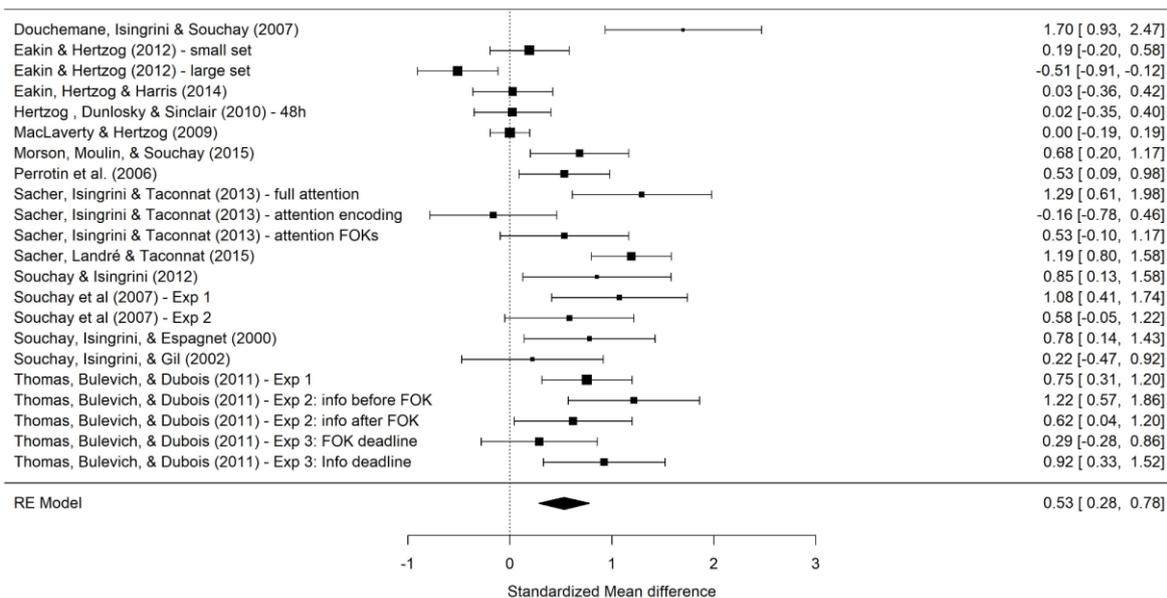
No study reporting sFOK showed a significant age effect on sFOK sensitivity (Experiment 1 in Allen-Burge & Storandt, 2000; Douchemane et al., 2007; Eakin et al., 2014; Marquié & Huet, 2000; Morson et al., 2015; Experiment 1 in Souchay et al., 2007), although memory performance was sometimes also unmatched as OA performed worse than YA (e.g., Douchemane et al., 2007; Souchay et al., 2007 when education is not controlled for) or even better than YA (Butterfield et al., 1988; Morson et al., 2015).

**Global meta-analyses**

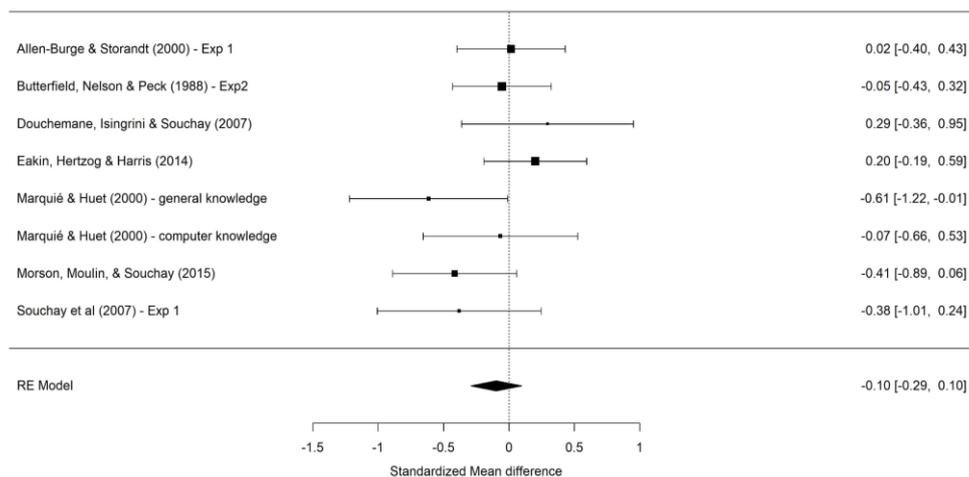
The M1e meta-analytic model showed lower gamma correlations in OA compared to YA,  $g = 0.53 [0.28, 0.78]$ ,  $p < .001$  (see Figure 4.2). The total heterogeneity analysis revealed a significant Q-statistic ( $Q(df = 21) = 102.77$ ,  $p < .001$ ). As we used a multilevel meta-analytic model, the amount of heterogeneity was computed for each level (Cheung, 2014) of the model (i.e., each random effect: studies, experiments, and effects).  $I^2$  for each level corresponds to 54.23%, 0%, and 24,86% respectively leading to 79.09% of the total variance due to heterogeneity.

The M1s model showed no age effect on gamma correlation across studies,  $g = -0.10 [-0.29, 0.10]$ ,  $p = .330$  (see Figure 4.3). The total heterogeneity analysis was non-significant,  $Q(df = 7) = 9.24$ ,  $p = 0.236$ . That is, we found no age difference in gamma correlations for the sFOK tasks.

According to our preregistration and because of significant heterogeneity in our eFOK model, we tested a moderation by memory performance. The M2 (recall model) revealed only a trend moderation effect,  $QM(df = 1) = 3.53$ ,  $p = .060$ . On the contrary, the M3 (recognition model) revealed a significant moderation effect,  $QM(df = 1) = 5.14$ ,  $p = .023$ . Moreover, for a recognition effect size of 0 (no group difference), the estimated effect does not reach significance (estimate = -0.24,  $p = 0.514$ ). Finally, the M4 model (type of recognition model) showed no effect of this moderator,  $QM(df = 1) = 0.52$ ,  $p = 0.469$ .

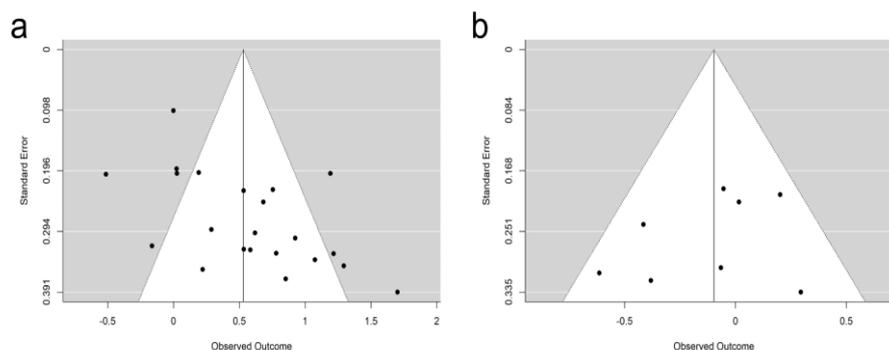


**Figure 4.2 :** Forrest plot of the effect of eFOK deficit in OA. Confidence interval of the overall estimated effect does not overlap with 0.



**Figure 4.3 :** Forrest plot showing an absence of sFOK deficit in OA. Confidence interval of the overall estimated effect does overlap with 0.

For the sFOK model, no publication bias was identified as the shape of the funnel plot showed no asymmetry ( $z = -0.91, p = .365$ ). However for the eFOK model, a significant asymmetry in the funnel plot was identified for eFOK: ( $z = 2.99, p = .003$ ) suggesting a publication bias (see Figure 4.4).



**Figure 4.4 :** Funnel plot centered on the overall effect size (vertical line) for eFOK model (a) and sFOK model (b).

The white areas are the 95% confidence intervals of the overall effect sizes. Points in the gray areas are outliers.

### 4.3.2. Analyses of individual scores from aggregated dataset

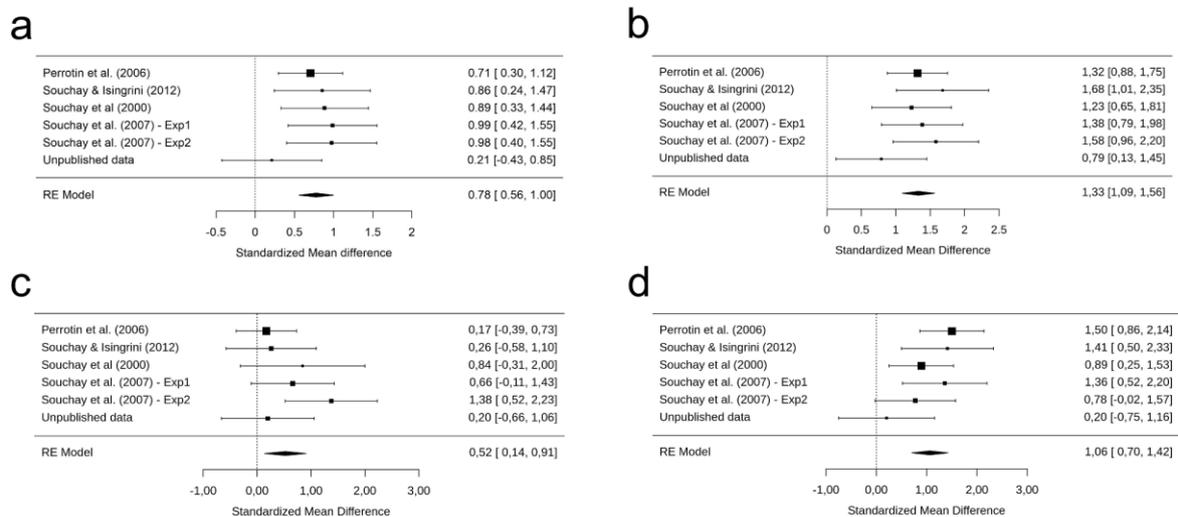
#### Qualitative review

In total, data from 235 older adults ( $M_{age} = 72.49; SD_{age} = 8.90$ ) and 134 younger adults ( $M_{age} = 24.91, SD_{age} = 3.14$ ) were analyzed. All studies used a standard eFOK paradigm split into 3 parts. Participants first studied pairs of cue-target words for 5 seconds. Following the learning phase, participants were presented with each cue and were given up to 15 seconds to recall the target associated with it. After each recall attempt, they made an FOK judgment.

They were asked to say “yes” when they thought they would be able to recognize the target later and “no” when they thought they would not be able to recognize it. Finally, participants were asked to identify the target among distractors. Four studies used a 5-alternative forced-choice paradigm in which each target was presented with 4 distractors and participants were asked to select the target. The two other studies presented a list of all targets together with the same number of distractors in which participants were asked to identify all targets. The six experiments data also slightly differed in the number of trials (36 or 40 word pairs).

**Meta-analyses**

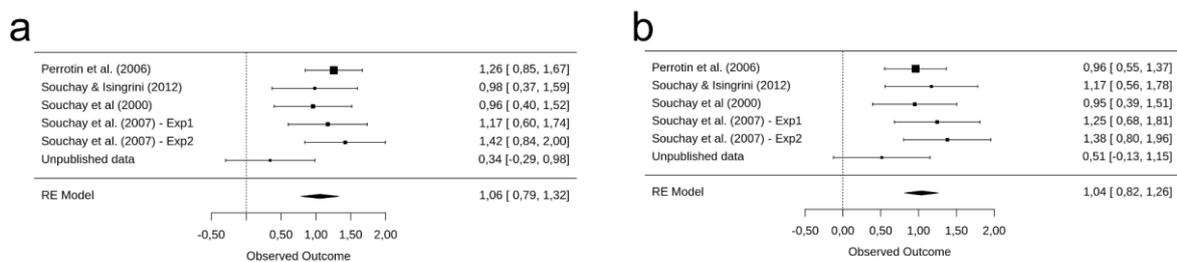
As for the M1e model, the M5 meta-analytic model showed lower gamma correlations in OA compared to YA,  $g = 0.77 [0.56, 0.99]$ ,  $p < .001$  (see Figure 4.5a). The total heterogeneity analysis revealed a non-significant Q-statistic ( $Q(df = 5) = 4.32$ ,  $p = .504$ ). The M6 model showed an overall large effect size of  $g = 1.33 [1.09, 1.56]$ ,  $p < .001$  (see Figure 4.5b) resulting in a higher performance in recall for YA compared to OA. Both M7a and M7b revealed a significant effect. Although the estimated effect size was lower for M7a,  $g = 0.52 [0.14, 0.91]$ ,  $p < .001$  than for M7b,  $g = 1.06 [0.70, 1.42]$ ,  $p < .001$  (see Figure 4.5c and Figure 4.5d), confidence intervals overlap. Note that these intervals are large as each model only includes half of the participants.



**Figure 4.5 :** Forrest plots of the different effects analyzed from aggregated dataset.

- a)** Forrest plot of the effect of eFOK deficit in OA. Confidence interval of the overall estimated effect does not overlap with 0.
- b)** Forrest plot of the effect of recall deficit in OA. Confidence interval of the overall estimated effect does not overlap with 0.
- c)** Forrest plot of the effect of eFOK deficit in OA for half of OA with the best recall performance and half of YA with the worst recall performance. Confidence interval of the overall estimated effect does not overlap with 0.
- d)** Forrest plot of the effect of eFOK deficit in OA for half of OA with the worst recall performance and half of YA with the best recall performance. Confidence interval of the overall estimated effect does not overlap with 0.

As exploratory analyses, we also conducted two other meta-analyses similar to M5. Instead of using corrected gamma as a measure of metacognitive sensitivity, we used the Hamann coefficient and type-II d'. Type-II d' was adjusted for extreme hits and false alarms by replacing rates of 0 with 0.5/n and rates of 1 with (n-0.5)/n where n is the number of signal or noise trials (Macmillan & Kaplan, 1985). As for the M5 model, the model with Hamann coefficient showed an overall effect of  $g = 1.06$  [0.79, 1.32],  $p < .001$  with no significant heterogeneity  $Q(df = 5) = 7.52$ ,  $p = .185$ , as well as the model with type-II d',  $g = 1.04$  [0.82, 1.26],  $p < .001$ ,  $Q(df = 5) = 4.88$ ,  $p = .431$  (see Figure 4.6a and Figure 4.6b).

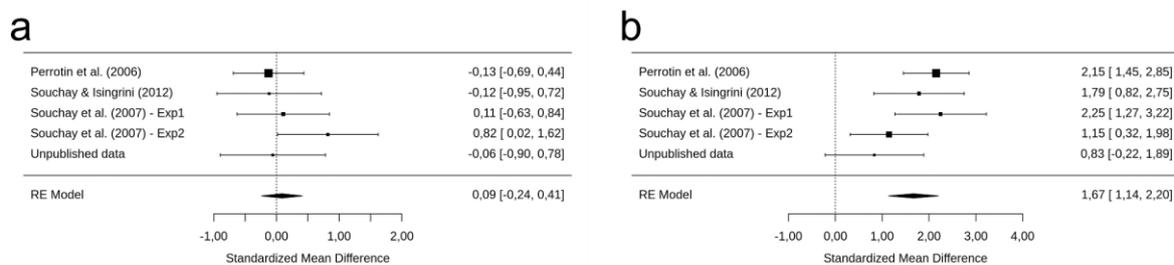


**Figure 4.6 :** Forrest plots of the effect of eFOK deficit in OA using other measures.  
 a) using the Hamann coefficient. Confidence interval of the overall estimated effect does not overlap with 0.  
 b) using type-II d'. Confidence interval of the overall estimated effect does not overlap with 0.

### Non-preregistered analyses

As the two meta-analyses splitting participants according to recall performance led to unclear results due to the small overlap between confidence intervals, we conducted a complementary analysis. The global meta-analysis revealed a moderator effect of recognition performance (and a trend for recall) going toward the idea that memory function is involved in eFOK sensitivity. This moderation by recognition however cannot disentangle between a proper memory-metamemory interaction, a spurious relationship due to gamma, or more likely both factors (because of the trend effect of recall). To see whether recognition performance better explains eFOK than recall also in our aggregated dataset, we performed the same split-analysis using recognition performance. The two datasets were therefore created according to the median split.

Akin to M7a and M7b, we created M8a and M8b. M8a showed a non-significant overall estimated effect,  $g = 0.09$  [-0.24, 0.41],  $p = .601$  whereas the effect for M8b was significant,  $g = 1.67$  [1.14, 2.20],  $p < .001$  (see Figure 4.7a and Figure 4.7b). Moreover, as confidence intervals do not overlap with each other, both models estimated different overall effects significantly.



**Figure 4.7 :** Forrest plots of of eFOK deficit in OA using median-splits based on recognition.

- a) For half of OA with the best recognition performance and half of YA with the worst recognition performance. Confidence interval of the overall estimated effect overlaps with 0.  
 b) For half of OA with the worst recognition performance and half of YA with the best recognition performance. The confidence interval of the overall estimated effect does not overlap with 0.

#### 4.4. Discussion du Chapitre 4

We conducted a systematic review and a meta-analysis of eFOK and sFOK in aging. We confirmed our main hypothesis of a preserved sFOK and a moderately impaired eFOK in aging ( $g = 0.53$ ). The large heterogeneity observed in the qualitative and the quantitative analyses is the result of a variety of methodologies in the different studies that were mostly used to control for memory differences between OA and YA. Such variability is not present in sFOK studies and OA have the same (or sometimes even better) performance in semantic memory tasks as YA. We also found that the aggregated dataset meta-analysis slightly inflated the estimation of the eFOK deficit in aging (from  $g = 0.53$  to  $g = 0.78$ ), probably due to the fact there was no attempt to control for first order performance (also confirmed by the non-significant heterogeneity for this meta-analytic model).

We investigated whether episodic memory performance explains the eFOK deficit observed in OA. Moderator analyses revealed that recall (as a trend) and especially recognition reduced the overall effect size. The analysis of the model's intercept suggests that when there is no difference between OA and YA in terms of recognition performance, the eFOK difference vanishes. Using aggregated dataset, reducing or increasing group difference in recall memory by selecting best/worst 50% of OA and worst/best 50% YA respectively only slightly reduced (from  $g = 0.78$  to  $g = 0.52$ ; see Figure 4.5c) or enlarged (from  $g = 0.78$  to  $g = 1.06$ ; see Figure 4.5d) the age-effects and the two models did not differ significantly. These analyses included only half of the sample and reduced statistical power could explain the overlap in confidence intervals of the effect sizes. Nonetheless, when performing the same median split analyses using recognition performance, modulations of the overall effect size were higher (from  $g = 0.78$  to  $g = 0.09$ , see Figure 4.7a; and from  $g = 0.78$  to  $g = 1.67$ , see Figure 4.7b). The model that decreased the recognition gap between groups led to a non-significant age-effect. As

such, when memory performance is equal across groups, the eFOK deficit in OA does not exist anymore.

Differences in recall and recognition therefore account for the different effects on eFOK sensitivity. This could be due to two reasons. First, we tackle contamination between measures of first order and second order performance, that is the age difference in accuracy could be a statistical artifact. Secondly, we discuss the possibility that there is a psychologically real reason why diminished memory performance leads to impoverished eFOK accuracy.

First, FOK sensitivity is calculated from recognition performance which explains that the larger the metacognitive sensitivity measure is biased by first order performance, the larger the estimated deficit. This is in line with our exploratory analysis showing a larger estimated effect using type-II  $d'$  than gamma correlation. Theoretically, type-II  $d'$  is less influenced by bias, as signal detection theory is proposed to distinguish between bias and sensitivity (but see Evans & Azzopardi, 2007). However, first order performance is more likely to influence this measure as type-II  $d'$  assumes that the distributions for “correct” and “incorrect” are Gaussian with equal variance which is rarely the case at the type-II level (Galvin et al., 2003). Metacognitive sensitivity measures have also been shown to be influenced by guessing during the recognition task in modeling on hypothetical data (Vuorre & Metcalfe, 2021). Gamma was drastically reduced as a function of guessing (i.e., for lower levels of knowledge) even when the relationship between first and second order performance was held constant in the model. The memory deficit in aging is variable but overall, we suggest that OA conform to this notion of middle-to-low performance as described by Vuorre and Metcalfe (2021).

Secondly, we propose that memory function is intrinsically linked to the capacity to make accurate metamemory judgements. If the eFOK difference in aging is due to an interaction between a metacognitive sensitivity measure and an episodic memory deficit, a pure memory hypothesis is likely also at play. FOKs are performed after recall. Thus, these judgments are undoubtedly directly influenced by the recall process (e.g., partial information available at this stage, Koriat, 1993; Schacter & Worling, 1985; Thomas et al., 2012). For example, Eakin & Hertzog (2012) showed that FOKs were more strongly correlated with recall than with recognition responses, both in YA and in OA. Moreover, in both eFOK and sFOK tasks, Mazancieux, Dinze, et al. (2020) showed that metacognitive efficiency (a measure of metacognitive sensitivity that controls for recognition) variability is more strongly correlated with recall than recognition. OA are proposed to have a trade-off between a deficit in the recollection process (also at play in recall tasks) and a preserved familiarity process (Parkin & Walter, 1992; Perfect et al., 1995). If the output of the retrieval process during recall leads to no information on which to base FOKs, OA are not able to anticipate recognition.

Importantly, we found that the eFOK deficit in aging is mainly explained by lower recognition in OA, and not recall. This points more in the direction of a contamination between

first and second order performance than a genuine problem of memory function, but this is something which needs further consideration. Therefore, we argue here for new studies that better control recognition performance between YA and OA. In the literature on confidence judgments, two main methods are used to achieve such control. The first is to use a metacognitive sensitivity measure that takes into account first order performance such as the meta- $d'/d'$  ratio (Fleming, 2017; Maniscalco & Lau, 2012, 2014) also known as metacognitive efficiency even though small dependencies between  $d'$  and the meta- $d'/d'$  ratio also exists particularly for low first order performance (Guggenmos, 2021).

The second method involves the experimental protocol that is used. Staircase procedures are often used to equate first order performance across groups or tasks although this has also been shown to inflate estimations of metacognitive efficiency (Rahnev & Fleming, 2019). Nonetheless, we propose that extensions of such protocols could be used in the FOK literature by for instance changing the distractors for trial  $n+1$  according to performance at trial  $n$ , or by manipulating other variables online which are critical for first order performance, such as study time, and retention interval.

Another simple solution would also be to measure episodic memory performance on a separate, perhaps standardized task, such that the episodic memory function and the FOK measure are not taken from the same task, thus sidestepping some of the issues of contamination between the measures statistically. Interestingly, if the eFOK was merely a statistical artifact or measurement issue, it would be expected that where OA semantic memory performance fell below that of YA people to a similar magnitude as episodic memory, we should also see parallels in sFOK performance: sFOK sensitivity should likewise be impaired where first order performance is deficient.

Finally, the eFOK meta-analysis funnel plot was asymmetric suggesting a publication bias. This is not surprising if the eFOK deficit in aging is mainly explained by a contamination between first (here the recognition task) and second order performance, considering that older adults are mainly impaired on recollection and can still solve recognition tasks based on familiarity (Parkin & Walter, 1992). According to task difficulty (e.g., number and type of distractors in the recognition task), substantial between-experiment variability can occur.

## CHAPITRE 4 : CONCLUSION

Cette méta-analyse met en évidence des différences liées à l'âge dans la sensibilité du eFOK. Ce déficit est clair à la fois en comparant les performances de eFOK des jeunes adultes et des adultes âgés, mais aussi en comparant les performances de sFOK et de eFOK chez les personnes âgées. Nous pensons que ce déficit est dû à (1) des effets confondus liés à l'interaction entre le calcul du score gamma, le *guessing* et les performances mnésiques plus faibles, et (2) des performances mnésiques inférieures résultant d'un déficit de recollection spécifique survenant au moment du rappel qui affecte la récupération des informations sur lesquelles les jugements métacognitifs se basent. Le facteur clé en jeu dans ces deux hypothèses concerne la performance de premier ordre, qui explique également le pattern de préservation du sFOK trouvé dans cet article (pour lequel il n'y avait pas de différence entre les groupes). Nous pensons que les futures études devraient prendre en compte cette variabilité de la performance de premier ordre entre les groupes afin d'étudier les aspects de la métacognition qui pourraient potentiellement différer entre YA et OA.

# Chapitre 5.

## Étude de la métacognition dans la démence sémantique

---

Comme nous l'avons vu dans le Chapitre 4, un déclin des compétences mnésiques peut s'accompagner d'une diminution de la capacité des personnes à évaluer l'exactitude de leur mémoire. Si la comparaison de jeunes adultes à des adultes âgés sur des tâches de FOK fournit un exemple de déclin de la sensibilité métacognitive lié à l'effet de l'âge sur la réduction des capacités en mémoire épisodique, dans ce chapitre, l'objectif était d'évaluer les capacités métacognitives de personnes présentant un déclin de la mémoire sémantique. Treize patients atteints de démence sémantique (DS) ont réalisé plusieurs tâches de métacognition : des tâches de PoR et de RCJ épisodiques et sémantiques. Leurs performances ont été comparées à celles de 20 participants sains et à celles de 7 patients atteints de maladie d'Alzheimer, une maladie neurodégénérative altérant principalement la mémoire épisodique. Cette étude avait pour objectif d'approfondir la compréhension de la relation entre type de déficit mnésique (i.e., épisodique ou sémantique) et compétences métacognitives. Nos résultats ne permettent pas de mettre en avant une quelconque diminution de la sensibilité métacognitive dans la démence sémantique. Nous discuterons des résultats en termes de relation entre performance de premier ordre et de sensibilité métacognitive, mais également de l'implication des deux formes de mémoire dans les jugements métacognitifs.

### 5.1. Introduction

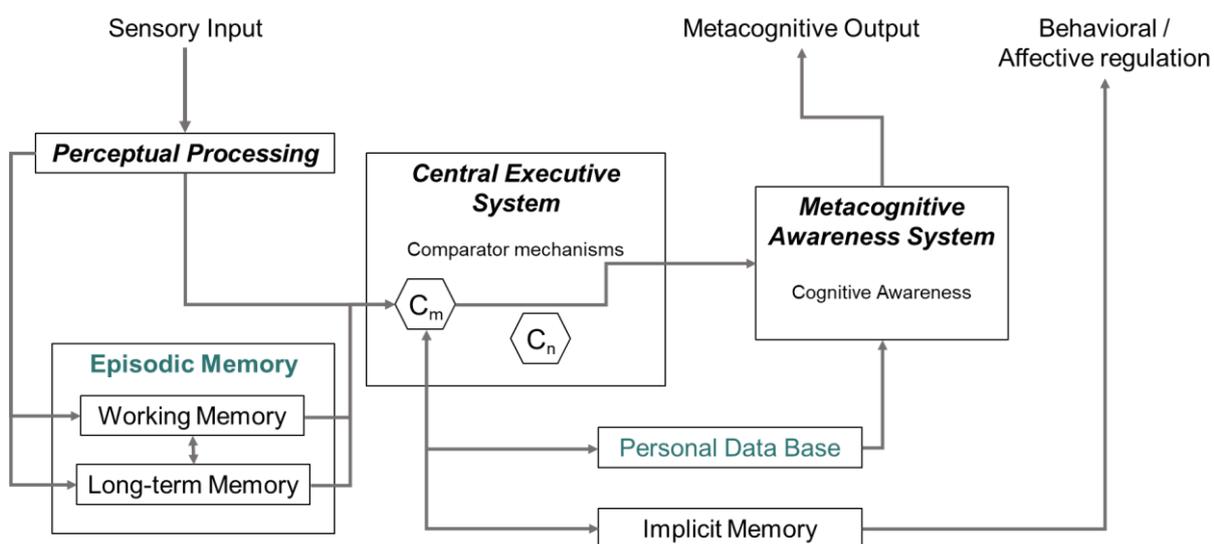
Les paradigmes de monitoring métacognitif fournissent des mesures objectives et des modèles explicatifs pour l'étude de la conscience qu'ont des patients atteints de troubles cognitifs de leur déficits (i.e., l'anosognosie). Ces paradigmes complètent les observations obtenues par d'autres méthodes, plus subjectives et très répandues dans le milieu clinique, comme les entretiens semi-structurés et des mesures de différence entre les évaluations des patients et celles de leurs proches (DeLozier & Davalos, 2016; Ruijter et al., 2020; Wilson et al., 2016). Les paradigmes de métacognition permettent une mesure objective de la capacité de personnes à auto-évaluer leurs performances sur différentes tâches cognitives. L'avantage est de pouvoir ensuite comparer leurs jugements à leur performance objective et de permettre

une analyse plus fine des facteurs en jeu dans leurs capacités à réfléchir sur leurs propres aptitudes. Selon un point de vue théorique, les pathologies de la mémoire fournissent un cadre pour étudier l'intervention et l'importance des processus mnésiques dans les tâches de métacognition.

Comme présenté en introduction de cette thèse, les études conduites sur la métacognition dans les pathologies de la mémoire ont majoritairement été réalisées auprès de populations de personnes présentant une dégradation de la mémoire épisodique (e.g., vieillissement, voir [Chapitre 4](#); maladie d'Alzheimer, MA, Ernst et al., 2015; Souchay, 2007; amnésie ; Janowsky et al., 1989; Shimamura & Squire, 1986). Dans les pathologies neurodégénératives de la mémoire, une grande majorité des travaux ont été réalisés auprès de patients MA (Bertrand et al., 2019; McGlynn & Kaszniak, 1991; Moulin et al., 2003; Souchay, 2007; Souchay et al., 2002, 2003 ; voir Ernst et al., 2015 pour une revue). Ces études ont montré que les patients MA ne présentaient pas un déficit métacognitif généralisé mais suggèrent plutôt un fractionnement de la métacognition dans cette pathologie (Ernst et al., 2015; Souchay, 2007). En particulier, les patients MA présentaient des difficultés se traduisant par une réduction de la sensibilité métacognitive sur les tâches de prédictions portant sur du matériel épisodique (e.g., eFOK ; Souchay et al., 2002). À l'inverse, des préservations de la sensibilité métacognitive ont été observées sur les tâches utilisant des RCJ (Moulin et al., 2003; Pappas et al., 1992), ou des FOK sémantiques (i.e., sFOK ; Bäckman & Lipinska, 1993; Lipinska & Bäckman, 1996 mais voir Pappas et al., 1992). Souchay (2007) a proposé que les difficultés observées sur les tâches de prédictions épisodiques chez les personnes MA résultaient d'une diminution des capacités de recollection, c'est-à-dire de la capacité à récupérer des détails relatifs à la phase d'apprentissage. Ces différentes études mettent en évidence un lien entre déclin de la mémoire épisodique et diminution de la sensibilité métacognitive sur des tâches portant sur ce type de mémoire.

D'un autre côté, les mémoires épisodique et autobiographique sémantique (i.e., connaissances reliées au soi), en bleu sur la Figure 5.1, sont mises au centre de certains modèles théoriques d'anosognosie qui tentent d'expliquer les processus permettant à une personne de pouvoir évaluer correctement ses aptitudes. D'après le modèle de *Cognitive Awareness* (CAM ; Agnew & Morris, 1998; Hannesdottir & Morris, 2007; Morris & Mograbi, 2013, voir Figure 5.1), lors d'un échec de récupération en mémoire (e.g., oubli d'un nom), la personne prend conscience de cet oubli grâce au système d'*awareness* métacognitive après un passage en mémoire de travail et par le système mnésique de comparaison (Cm). Le système de comparaison mnésique, contenu dans le système exécutif central, permet de comparer cet événement aux connaissances générales que la personne a de ses propres capacités mnésiques (i.e., croyances intra-personnelles, mémoire autobiographique sémantique), contenues dans la base de données personnelle (*personal database*). Si un

décalage est détecté entre cet évènement (i.e., l'oubli) et les connaissances personnelles du sujet (e.g., « je suis plutôt forte pour me souvenir des noms »), la base de données personnelle est mise à jour (e.g., « il m'arrive aussi parfois d'oublier des noms »). Le système épisodique permet de se souvenir des résultats des tentatives de récupération (i.e., souvenir ou oubli) précédentes. Ce modèle propose trois types d'anosognosies : mnémonique, exécutive et primaire. Dans sa forme mnémonique, l'anosognosie serait une conséquence d'un problème de mise à jour de la base de données personnelle en mémoire sémantique malgré une comparaison entre la mémoire épisodique et la base de données par le comparateur mnésique. Ce type d'anosognosie se manifesterait par un manque de conscience explicite du déficit (e.g., dans le discours) mais par des comportements implicites (e.g., évitement des situations mettant en jeu les compétences altérées). D'un autre côté, l'anosognosie exécutive résulterait d'un défaut du comparateur mnésique empêchant la comparaison de l'évènement d'oubli aux connaissances personnelles. Ainsi, les patients présentant ce type d'anosognosie percevraient leur erreur et en seraient conscients mais trouveraient des explications alternatives autres qu'un potentiel déficit en mémoire pour expliquer cet oubli. Enfin, dans l'anosognosie primaire, l'oubli n'est pas consciemment détecté et perçu comme étant une erreur, mais peut quand même être traité par la mémoire implicite.



**Figure 5.1** : Modèle CAM, reproduit d'après Morris & Hannesdottir (2004).

Ainsi, plusieurs travaux et modèles ont investigué la métacognition dans la MA et suggèrent l'intervention des processus épisodiques dans le déclin métacognitif, en particulier sur certaines tâches épisodiques dans cette pathologie. Cependant, encore peu d'études ont été conduites afin de voir si le lien entre déficit de premier ordre et sensibilité métacognitive

s'appliquait également sur des tâches sémantiques, et afin d'évaluer les performances métacognitives de personnes présentant des troubles plus sémantiques.

Cet aspect est d'autant plus important à considérer que les études conduites dans l'amnésie sans atrophie frontale ont montré qu'un déficit de premier ordre ne s'accompagnait pas systématiquement d'un déclin des capacités métacognitives sur les tâches mettant en jeu le type de mémoire affecté (Shimamura & Squire, 1986). Il est donc important de réaliser d'autres études afin d'investiguer plus profondément cette relation entre déficit mnésique et compétences métacognitives. De ce fait, l'étude de ce Chapitre 5 avait pour objectif d'évaluer les compétences métacognitives de personnes présentant un trouble neurodégénératif de la mémoire sémantique : la démence sémantique (DS).

La DS (Moreaud et al., 2008; Neary et al., 1998; Snowden et al., 1989) est une maladie neurodégénérative rare (touchant environ 3 personnes sur 100 000) qui apparaît en moyenne vers 60 ans (avant 65 ans dans 75% des cas, Belliard et al., 2010; Hodges et al., 2010) et qui affecte majoritairement la MS. Cette pathologie correspond à la forme temporelle des dégénérescences lobaires fronto-temporales (DFT ; mais voir Bonner et al., 2010; Gorno-Tempini et al., 2004; Grossman, 2010 pour des classifications de la DS comme la forme sémantique des aphasies primaires progressives). La DS correspond à environ 25% des cas de DFT (Belliard et al., 2010; Coyle-Gilchrist et al., 2016; Knopman & Roberts, 2011). Cette pathologie se caractérise notamment par la perte progressive du sens des mots et des concepts, et se manifeste majoritairement par des difficultés dans l'identification et la dénomination de personnes et d'objets, ainsi que par une altération de la compréhension des mots isolés (voir Hodges & Patterson, 2007 pour une revue des symptômes). Cette maladie évolue généralement de manière lente et progressive. Elle impacte dans un premier temps les connaissances anciennes et peu réactualisées, ainsi que les objets ou personnes peu familiers puis évolue vers une altération plus globale avec une réduction progressive du vocabulaire (Belliard et al., 2007). Cette pathologie résulte en particulier d'une atteinte cérébrale maximale au niveau des parties antérieures des lobes temporaux. Généralement, l'atteinte est à l'origine asymétrique (avec prédominance des atteintes gauches, Chan et al., 2001; Hodges et al., 1992) mais devient progressivement bilatérale avec l'avancée dans la maladie. Dans la DS, la composante épisodique de la mémoire est relativement préservée, notamment dans ses aspects non-verbaux, bien que des déficits de rappel soient également observés en mémoire épisodique sur les tâches utilisant du matériel verbal (Hodges et al., 1992; Hodges & Graham, 2001). Cette pathologie constitue donc un bon modèle pour investiguer la relation entre compétences mnésiques selon le type de mémoire testé (i.e., épisodique ou sémantique) et performances métacognitives.

À notre connaissance, la seule étude ayant porté sur la capacité des personnes DS à évaluer leurs performances cognitives suggère une anosognosie moins étendue que dans la forme comportementale de la DFT (vcDFT), et un niveau de *self-awareness* similaire à celui des patients MA (Eslinger, 2005). Dans cette étude, les patients réalisaient plusieurs tests cognitifs (fluence verbale, apprentissage d'une liste de mots). Pour chaque test, ils devaient rétrospectivement estimer la difficulté de la tâche (i.e., sorte de jugements globaux). Leurs jugements étaient ensuite corrélés, dans chaque groupe, à leur performance réelle. Des corrélations positives étaient observées entre les jugements et la performance que ce soit pour les contrôles, les patients DS ou les patients MA. De plus, les participants remplissaient également des questionnaires portant sur leurs comportements sociaux, cognitifs et émotionnels de la vie de tous les jours. Les évaluations des participants étaient ensuite comparées à celles rapportées par un de leurs proches (i.e., scores de différence, mesure subjective). Sur la majorité des sous-catégories de l'échelle, les scores de décalage n'étaient pas différents pour les patients DS, MA et contrôles. Cependant, les patients DS avaient tendance à sous-estimer la gravité de leurs symptômes apathiques. Ils surestimaient leur motivation quotidienne par rapport à ce qui était rapporté par leurs proches, ce qui était également le cas pour les MA. Les patients DS sous-estimaient également les changements comportementaux qu'ils manifestaient par rapport à leurs habitudes. Ainsi, cette étude montre que les patients DS ont globalement une certaine connaissance de leurs difficultés, malgré certains aspects comportementaux dont ils semblent avoir moins conscience.

Dans les DFT, plusieurs études portant sur les capacités d'auto-évaluation sur des tâches de mémoire ont été réalisées auprès de personnes atteintes de vcDFT (Banks & Weintraub, 2008; Bastin et al., 2012; Rosen et al., 2014; Souchay et al., 2003). La vcDFT se manifeste principalement par une désinhibition, de l'apathie, une perte de l'empathie, des comportements compulsifs, une hyperoralité, et un dysfonctionnement exécutif (Olney et al., 2017; Rascovsky et al., 2011). Ces études suggèrent un déficit métacognitif dans cette variante puisque les personnes vcDFT avaient notamment une tendance à surestimer leur mémoire épisodique (Banks & Weintraub, 2008), et avaient une moins bonne sensibilité métacognitive que des personnes âgées contrôles sur des tâches de eFOK et de eRCJ (Rosen et al., 2014, mais voir Bastin et al., 2012 pour la sensibilité du eFOK non significativement différente entre le vcDFT et les contrôles, bien que les scores gammas ne soient pas différents de zéro chez les vcDFT). Cependant, à notre connaissance, aucune étude n'a évalué les capacités métacognitives des personnes atteintes de DS sur des tâches de métacognition objectives, utilisant des jugements locaux et des mesures intra-individuelles plus spécifiques (e.g., sensibilité métacognitive, magnitude).

Ainsi, le but de cette expérience était de voir si un déficit métacognitif était observé dans une pathologie de la mémoire sémantique afin de fournir des résultats supplémentaires pour

mieux comprendre l'importance des capacités mnésiques sémantiques dans la métacognition. Pour cela, les capacités métacognitives de plusieurs patients atteints de DS, recrutés dans trois centres hospitaliers français, ont été comparées à celles de participants âgés sains appariés en âge sur des tâches épisodiques et sémantiques. Un second objectif était de mieux définir la relation entre type de déficit mnésique (i.e., épisodique ou sémantique) et profil de compétences métacognitives. Pour cela, les performances métacognitives des patients DS étaient comparées à celles de patients atteints de MA (i.e., trouble de la mémoire épisodique).

De précédents travaux suggèrent une plus grande corrélation entre capacités de rappel et efficacité métacognitive dans les tâches de PoR que dans les tâches de RCJ (Mazancieux, Dinze, et al., 2020), ce qui suggère que les processus mnésiques semblent être plus impliqués dans les jugements prédictifs que rétrospectifs. Ainsi, nous faisons donc l'hypothèse que la sensibilité métacognitive dépendrait plus des capacités mnésiques sur les tâches de PoR que sur des tâches de RCJ. De plus, et comme introduit dans le [Chapitre 2](#), les RCJ se basent sur des indices plus domaine-général (e.g., fluence de récupération) alors que les indices utilisés pour les PoR sont plus mémoire-spécifique (e.g., recollection). Les caractéristiques de la tâche en cours (e.g., le processus de remémoration sur une tâche de mémoire épisodique) peuvent être mises en jeu lors du jugement métacognitif, et la qualité de l'apprentissage d'une information est un facteur influençant la magnitude et la précision des eFOK (e.g., *memory constraint hypothesis* dans le vieillissement, Hertzog et al., 2010). Ainsi, sur la base de ces différents travaux et selon l'hypothèse d'un fractionnement de la métacognition (Souchay, 2007), nous nous attendions à ne pas observer les mêmes différences de sensibilité des jugements entre les groupes, selon le type de jugement métacognitif évalué (i.e., prospectif ou rétrospectif) et le type de mémoire testé (i.e., épisodique ou sémantique). Nous faisons l'hypothèse d'une différence de sensibilité métacognitive entre les patients DS et les contrôles. Les patients DS présentant une détérioration de la MS, nous nous attendions à ce que cette différence soit plus élevée lorsque l'item à rappeler mettait en jeu la mémoire sémantique que dans le cas d'une tâche épisodique, plus particulièrement sur les tâches de prédictions. Plus spécifiquement, nous prédisions que les patients DS auraient une moins bonne sensibilité métacognitive que les contrôles sur les tâches portant sur la mémoire sémantique, mais nous ne nous attendions pas à une telle différence sur les tâches portant sur la mémoire épisodique.

De plus, la mémoire épisodique étant plus touchée dans la MA que la DS, et la mémoire sémantique plus dans la DS que la MA (Simons et al., 2002), et en supposant un lien entre type de déficit de premier ordre et performance de second ordre, nous nous attendions à ce que la sensibilité métacognitive des personnes DS soit moins bonne que celle des MA sur des tâches sémantiques et meilleure que celle des MA sur des tâches épisodiques, en particulier sur les tâches de PoR (i.e., double dissociation).

## 5.2. Méthode

### 5.2.1. Participants

Treize patients ayant reçu le diagnostic de DS ont été inclus dans cette étude. Les patients ont été recrutés dans trois centres hospitaliers universitaires français (5 à Grenoble, 5 à Rennes et 3 à Angers). Leur atteinte cérébrale temporale était majoritairement à droite chez 5 patients, à gauche chez 4 patients et plutôt bilatérale chez 2 patients (voir Tableau 5.2). Ces données de latéralité étaient manquantes pour 2 patients. Six patients ayant reçu un diagnostic de MA ont également été recrutés au CHU-Grenoble Alpes, et un au CHU de Rennes. Le groupe contrôle était composé de 20 participants recrutés via la base de participants « Seniors pour la Science » (<https://seniorspourlascience.fr/>) de la Maison des Sciences de l'Homme de l'Université Grenoble Alpes, et était apparié au groupe DS en terme d'âge. Tous les participants contrôles avaient un MMSE supérieur au seuil pathologique de 27/30 (min = 27 ; max = 30). Aucun participant de cette étude ne présentait de trouble psychiatrique ou neurologique (autre que la DS ou la MA pour les patients), ou d'antécédents d'abus d'alcool ou d'autres drogues. Tous parlaient le français couramment. Les informations démographiques ainsi que les scores aux différentes échelles neuropsychologiques sont rapportés dans le Tableau 5.1. Le groupe contrôle était apparié au groupe DS en termes d'âge et de genre, mais le groupe contrôle avait un niveau d'éducation légèrement plus élevé que le groupe de patients DS, bien que cette différence ne fût pas significative. Des observations sur les patients DS sont précisées dans le Tableau 5.2. Ce projet a obtenu l'accord du Comité de Protection des Personnes (N°ID-RCB: 2020-A02236-33, projet META-DEM).

### 5.2.2. Protocole

Les séances commençaient par la passation du MMSE et de la BREF (si ceux-ci n'avaient pas déjà été réalisés dans les 4 mois précédents l'inclusion). Tous les participants réalisaient ensuite des tâches de métacognition permettant d'évaluer 3 types de jugements métacognitifs : jugements globaux<sup>2</sup>, prédictions de reconnaissance (PR) et jugements rétrospectifs de confiance (RCJ), voir Tableau 5.3 pour le récapitulatif. Chaque type de jugement était évalué sur une tâche de mémoire épisodique et une tâche de mémoire sémantique. Ainsi, les participants ont réalisé 6 tâches de métacognition sur le logiciel E-prime 3 (Psychology Software Tools, 2016). Les participants commençaient par les 2 tâches de jugements globaux, puis effectuaient les 2 tâches de PoR et terminaient par les 2 tâches de RCJ. Pour chaque type de jugement, l'ordre de passation entre la version épisodique et la version sémantique était contrebalancé entre les participants. Cependant, cet ordre restait le

---

<sup>2</sup> Nous avons décidé de ne pas inclure les résultats obtenus sur ces tâches car ils rentraient moins dans la problématique de la thèse. Ces tâches ne seront pas détaillées par la suite.

même pour toutes les types de jugements. Ainsi, si un participant commençait par la version épisodique des jugements globaux, il commençait également par la version épisodique des PoR et des RCJ. À la fin de ces tâches de métacognition, les participants remplissaient un questionnaire de plaintes mnésiques (McNair & Kahn, 1983) qui, lorsque possible, était également rempli par un proche. Dans ce cas, les proches devaient estimer les difficultés mnésiques des patients dans la vie de tous les jours, en répondant aux mêmes questions que les patients.

**Tableau 5.1** : Données démographiques des participants du projet META-DEM.

Deux tests de comparaisons ont été réalisés : un premier pour comparer les données des personnes DS aux participants contrôles ( $p_{DS/C}$ ) et un second pour comparer les personnes DS aux MA ( $p_{DS/MA}$ ). Les différences significatives (après correction de Bonferroni) sont soulignées.

	Contrôles (n= 20)	DS (n=13)	MA (n = 7)	$p_{DS/C}$	$p_{DS/MA}$
Age	69.60 (3.73)	71.46 (4.52)	77.43 (4.52)	.75	<u>0.02</u>
Hommes/Femmes	4/16	2/11	5/2	.74	<u>0.01</u>
Niveau d'éducation	14.75 (3.08)	11.5 (3.99)	14.86 (3.53)	.07	0.22
Anxiété (/21) <sup>a</sup>	6.35 (4.08)	7.54 (4.29)	8.29 (3.15)	1.00	1.00
Dépression (/21) <sup>a</sup>	4.10 (2.88)	6.38 (3.40)	6.00 (3.46)	.15	1.00
MMSE (/30) <sup>b</sup>	28.85 (0.81)	25 (2.75)*	24.71 (2.93)	<u>&lt;.001</u>	1.00
BREF (/18) <sup>c</sup>	17.10 (1.33)	14.45 (1.51)**	15.00 (1.63)	<u>&lt;.001</u>	1.00
Dénomination (/40) <sup>d</sup>		12 (8.21)***	37.57 (1.62)		<u>&lt;.001</u>
Appariements sémantiques visuels (/40) <sup>d</sup>		31.91 (5.39)**	38.86 (1.35)		<u>.002</u>
Appariements sémantiques verbaux (/40) <sup>d</sup>		30.18 (7.93)**	39.71 (0.49)		<u>&lt;.001</u>
Mémoire épisodique visuelle <sup>e</sup>					
Immédiate (/12)		7.22 (2.51)****	4.14 (2.23)		<u>.03</u>
Différée (/12)		6.39 (2.57)****	2.5 (1.83)		<u>.01</u>
Mémoire épisodique verbale <sup>e</sup>					
Immédiate (/12)		3.69 (2.70)*****	3.71 (2.38)		.95
Différée (/12)		3.19 (2.64)*****	2.43 (3.14)		.45
Echelle comportementale (/6)		3.33 (1.50)***	0.71 (0.76)		<u>.002</u>
(/40)		7.25 (4.71)***	0.86 (1.07)		<u>.002</u>

<sup>a</sup>Hospital Anxiety Depression Scale (Zigmond & Snaith, 1983); <sup>b</sup>Mini Mental State Examination (Folstein et al., 1975; Kalafat et al., 2003 pour la version française); <sup>c</sup> Batterie Rapide d'Efficiency Frontale (Dubois et al., 2000); <sup>d</sup>Batterie d'évaluation des connaissances sémantiques du GRECO, BECS-GRECO (Merck et al., 2011); <sup>e</sup>Batterie d'efficiency mnésique, BEM-84, (Signoret, 1991) ; <sup>f</sup>DAPHNE,(Boutoleau-Bretonnière et al., 2015).

\*Informations manquantes pour 3 patients; \*\* informations manquantes pour 2 patients; \*\*\* informations manquantes pour 1 patient; \*\*\*\* informations manquantes pour 4 patients; ; \*\*\*\*\* informations manquantes pour 5 patients

**Tableau 5.2 :** Descriptif des personnes DS : genre, âge, latéralité de l'atteinte, et observations particulières pendant la passation.

Participant	Genre	Age	Côté de prévalence de l'atrophie	Observations
1	Femme	70	D	Compréhension complexe des consignes. Ne comprend pas bien l'échelle de 0 à 10 sur les tâches de métacognition (e.g., dis « oui » au lieu d'un chiffre ou d'un niveau de certitude). Sur le tâche de PoR épisodique, semblait répondre plus sur sa capacité à se rappeler de l'item plutôt que de sa capacité future à reconnaître l'item parmi des propositions.
2	Homme	80	G	Bonne compréhension des consignes.
3	Femme	66	G	Bonne compréhension des consignes. Semblait comprendre l'échelle de jugements mais n'arrivait parfois pas à répondre.
4	Femme	73	Bilatérale	Bonne compréhension des consignes.
5	Femme	73	D	Bonne compréhension des consignes.
6	Femme	71		Compréhension complexe des mots isolés mais semblait avoir une bonne compréhension des consignes.
7	Homme	63		Bonne compréhension des consignes.
8	Femme	71	D	Bonne compréhension des consignes.
9	Femme	75	D	Compréhension complexe des consignes. Quelques incohérences entre ses dires (e.g., « je ne sais pas ») et ses réponses (e.g., appuie sur « 10, très certaine »).
10	Femme	72	G	Bonne compréhension des consignes.
11	Femme	77	Bilatérale	Compréhension complexe des consignes.
12	Femme	71	D	Semblait bien comprendre les consignes. Quelques incohérences entre ses dires (e.g., dit qu'elle répond au hasard) et ses réponses (e.g., appuie sur « 10, très certaine »).
13	Femme	67	G	Compréhension complexe des consignes.

D = droite, G = gauche ; les cases non remplies signifient que les données n'ont pas été renseignées.

### Tâche de prédiction de reconnaissance (PoR) sémantique.

Dans un premier temps, les participants voyaient 20 noms d'objets (e.g., stroboscope, voir [Annexe B](#) pour la liste) et devaient indiquer sur une échelle en 6 points allant de 0 (*Pas du tout certain*) à 10 (*Très certain*) à quel point ils étaient certains de pouvoir reconnaître l'image correspondant à ce mot (i.e., PoR). Ensuite, ils devaient identifier les images correspondant à chaque mot parmi deux propositions (voir Figure 5.2).

### Tâche de PoR épisodique.

Dans une première phase, les participants apprenaient 20 paires indice-cible d'images (e.g., un cheval à gauche et une banane à droite). Les paires étaient présentées pendant 1,5s (2s pour les MA, version facile)<sup>3</sup> et séparées par un intervalle de 250ms (500ms pour les MA). Ensuite, les 20 images vues à gauche (i.e., indices) étaient présentées seules et pour chacune, tous les participants devaient prédire sur la même échelle de 0 à 10 à quel point ils étaient certains qu'ils pourraient reconnaître la cible (i.e., image qui était vue avec lors de la phase d'apprentissage). Enfin, les participants voyaient les images indices à nouveau devaient

<sup>3</sup> Les personnes MA ayant plus de difficultés en mémoire épisodique que les personnes DS (Simons et al., 2002), le temps de présentation pour ces personnes était plus long. Le but était de faciliter l'encodage des items et ainsi tenter de réduire la différence de performance de premier ordre entre ces personnes et les personnes DS.

identifier la cible parmi deux images. Afin d'éviter une performance plafond et d'augmenter le temps entre l'apprentissage et les phases de prédiction et de test, nous avons ajouté une tâche de distraction pour les participants contrôles (version difficile). Dans cette tâche, les participants voyaient des images d'items insérés dans des scènes sémantiquement congruentes ou incongruentes (images des Expériences 7.1 et 7.2). Pour chacune, ils devaient indiquer si l'objet était congruent ou non avec la scène (voir Figure 5.2).

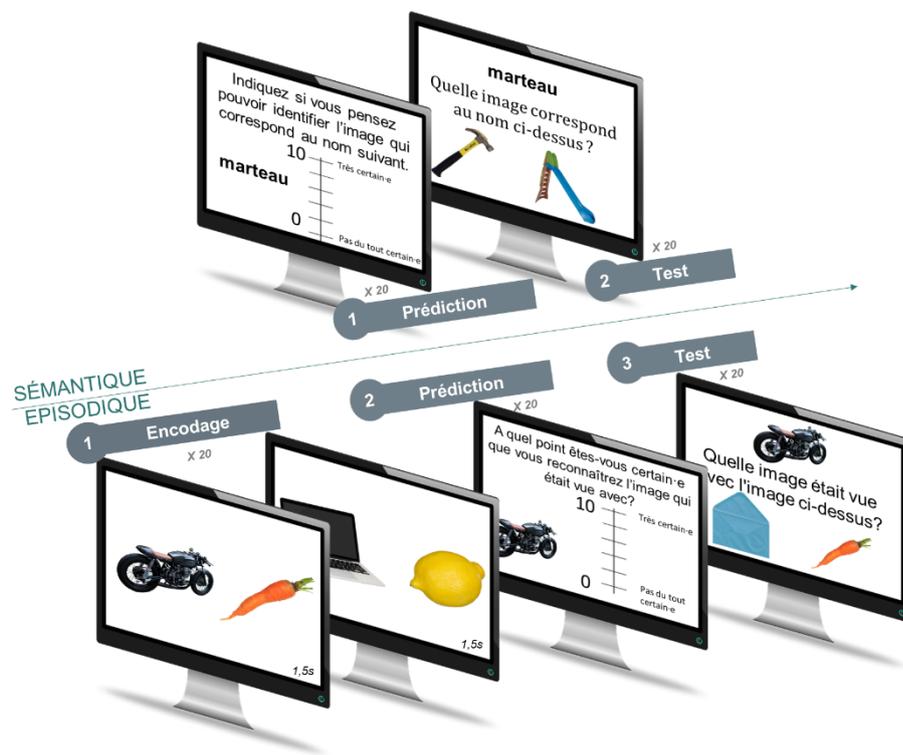


Figure 5.2 : Procédures utilisées pour les tâches de PoR.

### Tâche de RCJ sémantique.

Pour chaque essai, un mot-clé faisant référence à une profession, une nationalité, ou un sport (e.g., président, anglais, golf) était présenté avec deux photos de personnes célèbres. Les participants devaient indiquer à laquelle des deux personnes le mot-clé correspondait le plus (e.g., Emmanuel Macron ou Jean-Pierre Foucault). Après chaque réponse, les participants devaient juger l'exactitude de leur réponse sur une échelle de 0 (*pas du tout certain*) à 10 (*très certain*) (voir Figure 5.3).

### Tâche de RCJ épisodique.

Dans une première phase, les participants devaient mémoriser 20 visages humains présentés pendant 1,5s (2s pour les MA, version facile) et séparés des autres par 250ms (500 ms pour

les MA). Dans la phase de test, chacun des 20 visages était présenté avec un autre visage distracteur jamais vu et à chaque essai, les participants devaient d'abord identifier la cible et indiquer sur l'échelle de 0 à 10 à quel point ils étaient certains d'avoir sélectionné le bon visage. Pour les participants contrôles, la même tâche de distraction que pour la tâche de PoR épisodique a été ajoutée entre l'encodage et la phase de test/RCJ mais d'autres images étaient présentées (version difficile) (voir Figure 5.3).

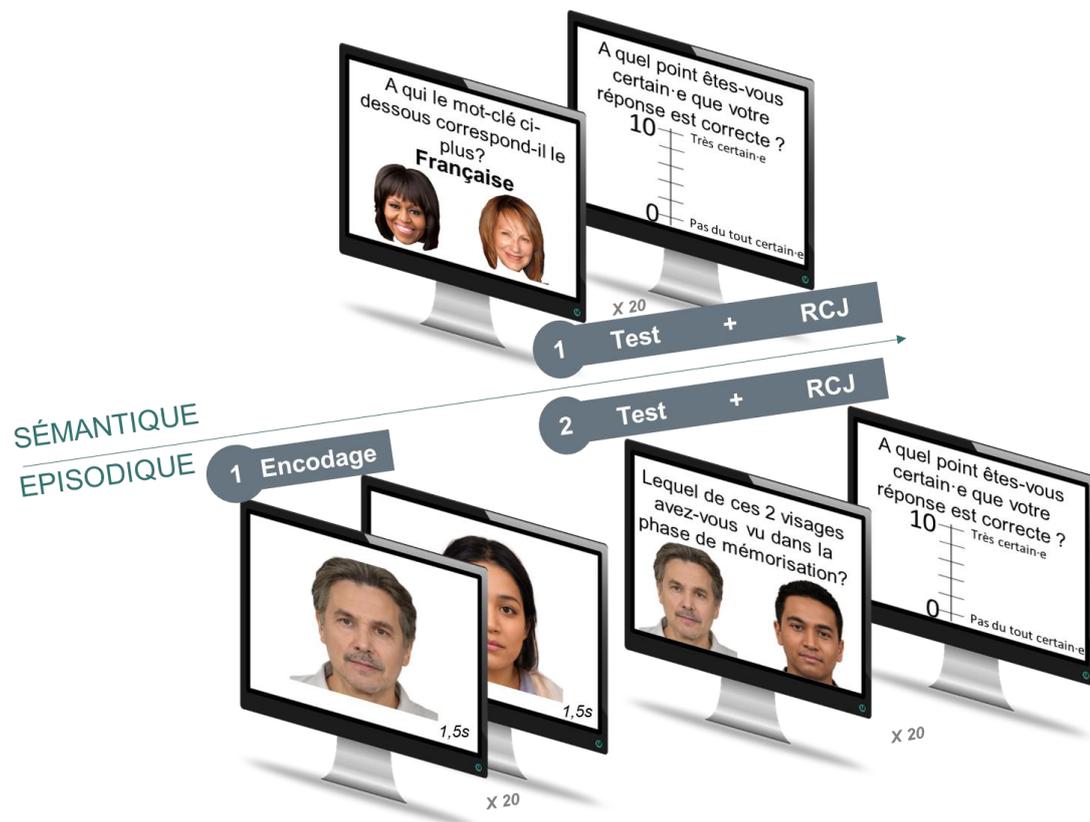


Figure 5.3 : Procédures utilisées pour les tâches de RCJ.

### 5.2.3. Matériel

Plusieurs études précédentes ont montré que les capacités métacognitives pouvaient être impactées par la difficulté de la tâche de premier ordre (Galvin et al., 2003; Masson & Rotello, 2009; Vuorre & Metcalfe, 2021). Ainsi, afin de pouvoir tirer des conclusions sur les capacités métacognitives plus « pures », nous avons créé des versions différentes de chaque tâche pour essayer de réduire les différences de performances de premier ordre entre les groupes (e.g., augmentation du temps de présentation ou introduction d'une tâche distractive sur les tâches épisodiques, voir ci-dessous pour les tâches sémantiques). En raison des difficultés de compréhension de mots observées chez les personnes DS (Belliard et al., 2011; Neary et al., 1998), afin que les participants comprennent bien que les consignes changeaient

d'une tâche à l'autre (e.g., consignes différentes entre les tâches de PoR et de RCJ), et puisque l'intérêt n'était pas de comparer directement les deux types de jugements, plusieurs stimuli ont été utilisés : des objets pour les tâches de PoR et des visages pour les tâches de RCJ. Des études ont montré que les personnes DS pouvaient montrer des difficultés sur des tâches de mémoire épisodique lorsque du matériel verbal était utilisé (e.g., Hodges & Graham, 2001), des images ont donc été utilisées. Afin d'éviter une contamination du déficit sémantique sur les tâches épisodiques, des objets très fréquents (i.e., identifiés par les personnes DS)<sup>4</sup> ont été choisis pour la tâche de PoR épisodique. Pour la tâche de eRCJ, des visages de personnes qui n'existent pas, générés par une IA ont été sélectionnés (<https://generated.photos/faces>).

Pour les tâches sémantiques, les participants contrôles, DS, et MA ont réalisé des versions différentes des tâches PoR et RCJ afin d'éviter que ces tâches ne soient trop complexes pour les personnes DS (en raison de leur trouble sémantique) ni trop simples pour les participants contrôles et MA, et ainsi limiter le risque de ne pas pouvoir calculer de scores métacognitifs (e.g., effets plafonds ou manque de variation des jugements métacognitifs si trop complexe), et afin d'essayer d'avoir des performances de premier ordre équivalentes dans ces 3 groupes. Pour cela, les différentes versions des tâches de PoR et de RCJ sémantiques de niveaux de difficulté différents ont été conçues sur la base d'un pré-test réalisé en ligne (i.e., version simplifiée et version complexe ; voir [Annexe B](#) pour le matériel utilisé pour la tâche de PoR sémantique, voir **Encart 1** ci-dessous). Chacune des versions de la tâche de sRCJ comprenait autant de visages féminins que masculins. Les distracteurs de la tâche de reconnaissance étaient du même genre que la cible. Certaines des personnes représentées étaient décédées et d'autres encore en vie.

Pour les tâches sémantiques, tous les contrôles ont passé les versions difficiles. Cependant, plusieurs versions ont pu être passées par les personnes DS. En effet, à part les 3 premiers patients rencontrés qui ont uniquement passé la version facile de la tâche de PoR, tous les autres patients ont passé la version intermédiaire. Ensuite, ceux pour qui cette version était trop compliquée (i.e., ils rapportaient répondre au hasard tout le long) ont également réalisé la version facile. De même, sur la tâche de RCJ, la version facile était administrée en premier aux patients, puis dans le cas où les performances étaient trop élevées (100%RC), la version difficile était ensuite administrée. Pour ces participants ayant réalisé plusieurs versions d'une même tâche, les données de la version dans laquelle les participants montraient des performances moins extrêmes (i.e.,  $0.55 < \%RC < 0.95$ ) ont été gardées pour les analyses. Dans les cas où les performances étaient extrêmes sur les deux versions (en général, plafond

---

<sup>4</sup> Pour 10 patients DS, il a été vérifié que les items étaient identifiés (dénomination ou description de l'utilité ou du contexte d'utilisation). Tous l'étaient pour la plupart, sauf deux items peu reconnus (gâteau et collier).

sur l'une et plancher sur l'autre), les données de la version dans laquelle la valeur absolue de la note Z pour la performance de premier ordre était la plus faible étaient conservées.

### Encart 1. Description de l'étude pilote pour la sélection des stimuli des tâches sémantiques.

Dans un pré-test, 45 personnes sans pathologie neurologique ou psychiatrique âgées entre 56 et 88 ans ( $M_{age} = 65.58$ ,  $SD_{age} = 5.87$ , 28 femmes) ont réalisé deux tâches sémantiques de l'étude finale : une de PoR et l'autre de RCJ. Les procédures étaient identiques à celles utilisées dans l'étude finale (décrites dans la section Protocole) mais le nombre d'items sur lesquels les participants étaient testés était plus conséquent et les prédictions étaient faites sur une échelle en 5 points.

Pour la tâche de PoR, les participants ont été testés sur leur capacité à reconnaître parmi deux propositions les images correspondant à 84 mots. Nous étions intéressés par leurs prédictions et leurs performances réelles. Sur la base des résultats, nous avons initialement créé une version simplifiée de la tâche de PoR sémantique pour les personnes DS avec les essais les plus réussis avec des prédictions élevées (e.g., saxophone, trampoline) et une version plus complexe pour les participants contrôles et MA avec les items les moins bien réussis et des prédictions variables (e.g., grelinette, piccolo). En choisissant les items les plus réussis pour la tâche simplifiée (version facile), nous pensions vérifier qu'il s'agissait d'items normalement connus par les personnes du même âge, et supposons donc que les personnes DS les avaient connus à un moment de leur vie. De plus, en discussion avec les neurologues, neuropsychologues, et orthophonistes, nous pensions en revanche que beaucoup de ces items pourraient ne pas être reconnus par les personnes DS. Cependant, les 3 premières personnes DS rencontrées ayant montré des performances plafonds pour cette tâche (entre 90 et 100%RC, avec une majorité de prédictions « *Très certain-e de le reconnaître* »), d'autres items moins fréquents (Lexique, New et al., 2004) non pré-testés (e.g., stéthoscope, caribou) ont finalement été sélectionnés dans une troisième version de la tâche (version intermédiaire, simplifiée mais plus compliquée que la version facile), utilisée pour les personnes DS vues par la suite. Les images étaient présentées en couleur et représentaient des animaux, des aliments, ou des objets.

Dans la tâche de RCJ, comme dans la tâche finale, les participants devaient reconnaître la personne à laquelle le mot-clé faisait référence et indiquer leur degré de certitude en leur réponse. Ils étaient testés sur 90 essais. Le trouble d'identification des visages étant caractéristique de la démence sémantique (Snowden et al., 2004), les items les mieux réussis par les participants sains du pré-test ont été sélectionnés pour créer la version simplifiée de la tâche. Les items les moins bien réussis ayant reçu des jugements de confiance divers ont été sélectionnés pour la version complexe de la tâche.

**Tableau 5.3** : Récapitulatif des tâches de métacognition et du matériel utilisé.

	<b>Prédictions de reconnaissance (PoR)</b>	<b>Jugements rétrospectifs de confiance (RCJ)</b>
<b>Sémantique</b>	20 images correspondant à des noms d'objet/animaux/ aliments	20 visages de personnes célèbres en lien avec mot-clé
<b>Épisodique</b>	20 paires d'images d'objet/animaux/ aliments	20 visages générés artificiellement

#### 5.2.4. Analyses

Des analyses par groupe ont été conduites pour évaluer l'effet du groupe (Contrôles, DS, MA) et du type de mémoire (épisodique, sémantique) sur les performances de premier ordre, sur la magnitude des jugements et sur la sensibilité métacognitive mesurée avec l'AUROC2. Ces analyses ont été faites pour les tâches de PoR puis pour les tâches de RCJ. En raison d'une erreur de programmation, les données de 4 personnes (3 DS, 1 MA) n'ont pas été récoltées pour la tâche de RCJ épisodique. Les données de ces 4 participants ont donc été exclues de toutes les analyses conduites à propos de la tâche de RCJ. Les conditions d'application (normalité des résidus ou homoscedasticité) n'étant pas respectée sur plusieurs modèles utilisés et pour plus de cohérence, nous avons utilisé des tests non-paramétriques pour l'ensemble des analyses (i.e., Kruskal-Wallis pour l'effet du groupe et Wilcoxon pour l'effet du type de mémoire). Pour les différentes analyses portant sur la magnitude, par soucis de comparaison avec la performance réelle, les jugements étaient transformés sur une échelle de 0.5 à 1 (0 = 0.5 ; 10 = 1). Pour les tâches de PoR et RCJ, des analyses complémentaires, plus individuelles et descriptives ont également été réalisées. Les moyennes et écart-types de chacune des tâches sont disponibles en [Annexe C](#).

Pour les questionnaires de plaintes mnésiques, deux ANOVA ont été conduites : une première pour évaluer la différence entre les scores de patients et de leur proche selon les groupes, et une pour calculer le score de plaintes des patients selon le groupe.

### 5.3. Résultats

#### 5.3.1. Tâches de PoR

##### **Performance mnésique (%RC)**

Le %RC était plus élevé sur la tâche épisodique ( $M = 0.83$ ,  $SD = 0.15$ ) que sur la tâche sémantique ( $M = 0.76$ ,  $SD = 0.10$ ) ( $V = 483$ ,  $p = .02$ ) mais ne différait pas significativement entre les 3 groupes ( $\chi^2 = 4.72$ ,  $p = .09$ , voir Figure 5.4). Cependant, l'effet du groupe dépendait du type de mémoire testé ( $\chi^2 = 14.23$ ,  $p = .001$ ,  $\eta^2 = 0.33$ ). En effet, dans la tâche épisodique, le %RC des personnes DS était plus faible ( $M = .80$ ,  $SD = .11$ ) que celui des contrôles ( $M = 0.90$ ,  $SD = 0.08$ ) ( $V = 197.5$ ,  $p_{\text{bonferroni}} = .037$ ). La différence entre les personnes DS et MA ( $M = 0.69$ ,  $SD = .24$ ) n'était pas significative ( $V = 57.5$ ,  $p_{\text{bonferroni}} = 1.00$ ). À l'inverse, sur la tâche sémantique, les participants DS ( $M = 0.83$ ,  $SD = 0.12$ ) ont eu de meilleures performances que les contrôles ( $M = 0.72$ ,  $SD = 0.09$ ) ( $V = 52.5$ ,  $p_{\text{bonferroni}} = .012$ ), et que les MA ( $M = 0.74$ ,  $SD = 0.02$ ) ( $V = 77$ ,  $p_{\text{bonferroni}} = .038$ ). La différence entre les contrôles et les MA n'était pas significative, que ce soit sur la tâche épisodique ( $V = 106$ ,  $p_{\text{bonferroni}} = .14$ ) ou sémantique ( $V = 61$ ,  $p_{\text{bonferroni}} = 1.00$ ).

**Magnitude**

La magnitude des PoR n'était globalement pas différente selon la mémoire testée ( $V = 322$ ,  $p = .35$ ), ni le groupe ( $\chi^2 = 1.42$ ,  $p = .49$ ). L'interaction entre le type de mémoire et le groupe n'était pas significative ( $\chi^2 = 1.75$ ,  $p = .42$ , voir Figure 5.4).

Pour compléter, et donner une première idée de la capacité des personnes à discriminer les essais corrects des essais incorrects, nous avons analysé la magnitude des jugements, en fonction de la performance de premier ordre. Ces analyses ont révélé que la magnitude des PoR était plus élevée pour les essais ensuite réussis ( $M = 0.71$ ,  $SD = 0.13$ ) que pour les essais échoués ( $M = 0.64$ ,  $SD = 0.12$ ) ( $V = 559$ ,  $p < .001$ ). Cependant, la différence entre la magnitude des essais incorrects et des essais corrects était plus grande sur la tâche sémantique ( $V = 653$ ,  $p < .001$ ) que sur la tâche épisodique ( $V = 466.5$ ,  $p = .013$ ). Il n'y avait pas d'interaction entre l'effet de la performance et le groupe ( $\chi^2 = 2.52$ ,  $p = .28$ ). La triple interaction performance\*mémoire\*groupe n'était pas non plus significative ( $\chi^2 = 0.10$ ,  $p = .95$ ).

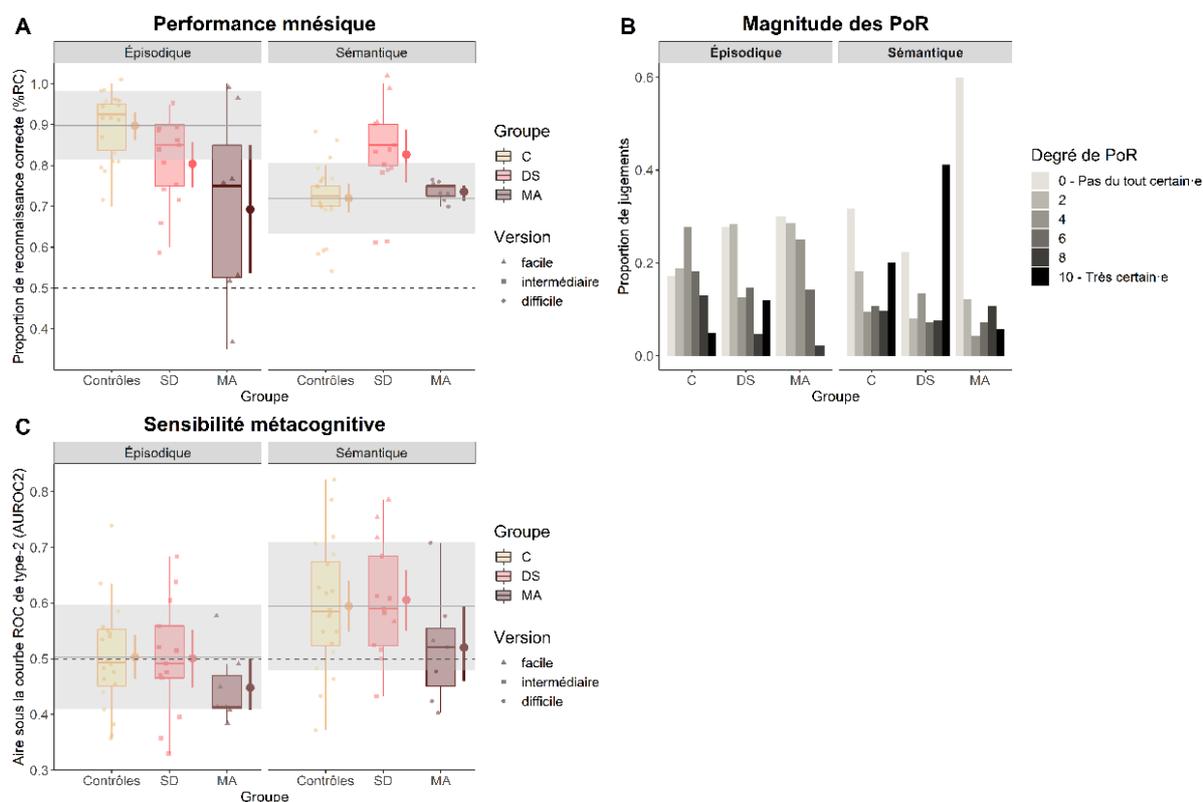
**AUROC2 (sensibilité métacognitive)**

Les scores AUROC2 n'étaient pas différents selon le groupe ( $\chi^2 = 4.08$ ,  $p = .13$ ) mais étaient plus élevés sur la tâche sémantique ( $M = 0.59$ ,  $SD = 0.11$ ) qu'épisodique ( $M = 0.49$ ,  $SD = 0.09$ ) ( $V = 138$ ,  $p < .001$ ), voir Figure 5.4. L'effet du groupe ne différait pas significativement selon la forme de mémoire testé ( $\chi^2 = 0.40$ ,  $p = .82$ ). Dans la tâche épisodique, les scores AUROC2 n'étaient significativement différents de 0.5 (i.e., PoR au hasard) dans aucun des groupes (DS :  $M = 0.50$ ,  $SD = 0.10$ ,  $V = 43$ ,  $p = .89$  ; MA :  $M = 0.45$ ,  $SD = 0.07$ ,  $V = 3$ ,  $p = .08$ ; contrôles :  $M = 0.50$ ,  $SD = 0.09$ ,  $V = 104$ ,  $p = .99$ ). Dans la tâche sémantique, les PoR des personnes DS ( $M = 0.61$ ,  $SD = 0.10$  ;  $V = 74$ ,  $p = .007$ ) et des contrôles était significativement différente du hasard ( $M = 0.59$ ,  $SD = 0.21$  ;  $V = 183$ ,  $p = .004$ ) mais pas celles des personnes MA ( $M = 0.52$ ,  $SD = 0.10$  ;  $V = 15$ ,  $p = .94$ ).

**Contrôle de l'âge**

Puisque les personnes MA étaient plus âgées que les contrôles et les DS, nous avons analysé l'effet de l'âge sur les différentes mesures étudiées. L'âge corrélait significativement uniquement avec la performance mnésique globale ( $r = -0.60$ ,  $p < .001$ ), et plus particulièrement avec la performance mnésique sur la tâche épisodique ( $r = -0.68$ ,  $p < .001$ ). Nous avons donc réalisé une ANCOVA afin de contrôler l'âge dans l'analyse de l'effet du groupe sur les performances de premier ordre. Après contrôle de l'âge, l'interaction entre le type de mémoire et le groupe restait significative ( $F(2,36) = 5.50$ ,  $p = .008$ ,  $\eta_p^2 = 0.23$ ). Sur la tâche épisodique, la différence entre le %RC des contrôles et des personnes DS n'était plus significative ( $t(36) = 1.63$ ,  $p_{\text{bonferroni}} = 0.11$ ). En revanche, sur la tâche sémantique, la différence entre les personnes DS et les contrôles restait significative ( $t(36) = -3.23$ ,  $p_{\text{bonferroni}} = 0.003$ ),

mais celle entre les personnes DS et MA ne l'était plus ( $t(36) = 1.62$ ,  $p_{\text{bonferroni}} = 0.11$ ). Les autres comparaisons ne changeaient pas après contrôle de l'âge ( $ps > .05$ ).



**Figure 5.4 :** Résultats sur la tâche de PoR.

A) Performance de premier ordre, B) Magnitude des RCJ, C) AUROC2. Les points épais sont les moyennes par condition, avec l'intervalle de confiance à 95%. Les points plus petits sont les données pour chaque participant. Les zones grises indiquent les scores situés à moins d'un écart-type de la moyenne des contrôles.

### 5.3.2. Tâches de RCJ

#### **Performance mnésique (%RC)**

La performance de premier ordre différait entre les groupes ( $\chi^2 = 14.40$ ,  $p < .001$ ). En effet, les personnes DS ont reconnu moins de cibles ( $M = 0.68$ ,  $SD = 0.14$ ) que les contrôles ( $M = 0.82$ ,  $SD = 0.13$ ) ( $V = 179$ ,  $p_{\text{bonferroni}} = .002$ ), voir Figure 5.5. Le %RC des personnes MA ( $M = 0.70$ ,  $SD = 0.16$ ) ne différait significativement pas de celui des personnes DS ( $V = 23.5$ ,  $p_{\text{bonferroni}} = 1.00$ ), ni des contrôles ( $V = 98$ ,  $p_{\text{bonferroni}} = .07$ ). Le %RC n'était pas significativement différent selon le type de mémoire testé ( $V = 337$ ,  $p = .17$ ). L'interaction entre le type de mémoire et le groupe n'était pas significative ( $\chi^2 = 2.68$ ,  $p = .26$ ).

**Magnitude des RCJ**

La magnitude moyenne des RCJ n'était pas différente selon le type de mémoire testé ( $V = 323.5$ ,  $p = .14$ ) ni selon le groupe ( $\chi^2 = 4.54$ ,  $p = .10$ ). L'effet du groupe n'interagissait pas avec le type de mémoire testé ( $\chi^2 = 2.15$ ,  $p = .34$ , voir Figure 5.5).

En prenant en compte la performance de premier ordre, les analyses ont montré que la magnitude des RCJ était plus élevée pour les essais corrects ( $M = 0.83$ ,  $SD = 0.12$ ) que pour les essais incorrects ( $M = 0.78$ ,  $SD = 0.14$ ) ( $V = 426$ ,  $p < .001$ ). Cet effet n'était pas significativement différent selon le type de mémoire testé ( $V = 301$ ,  $p = .30$ ) et ne différait pas entre les groupes ( $\chi^2 = 0.94$ ,  $p = .62$ ). L'interaction entre la performance et le groupe ne dépendait pas du type de mémoire testé ( $\chi^2 = 2.32$ ,  $p = .31$ ).

**AUROC2 (sensibilité métacognitive)**

Les scores AUROC2 étaient différents entre les groupes ( $\chi^2 = 8.35$ ,  $p = .02$ , voir Figure 5.5). Les scores AUROC2 des personnes DS ( $M = 0.58$ ,  $SD = 0.11$ ) et des MA ( $M = 0.55$ ,  $SD = 0.12$ ) étaient plus faibles que ceux des contrôles ( $M = 0.65$ ,  $SD = 0.12$ ) bien que ces comparaisons deux à deux ne soient que tendanciellées après correction de Bonferroni (DS *versus* contrôles :  $V = 152$ ,  $p_{\text{bonferroni}} = .065$  ; MA *versus* contrôles :  $V = 97$ ,  $p_{\text{bonferroni}} = .07$ ). La différence entre les personnes DS et MA n'était pas significative ( $V = 38$ ,  $p_{\text{bonferroni}} = 1.00$ ). Les scores AUROC2 n'étaient pas différents selon le type de mémoire testé ( $V = 362$ ,  $p = .66$ ). L'interaction entre l'effet du groupe et le type de mémoire n'était pas significative ( $\chi^2 = 3.78$ ,  $p = .15$ ).

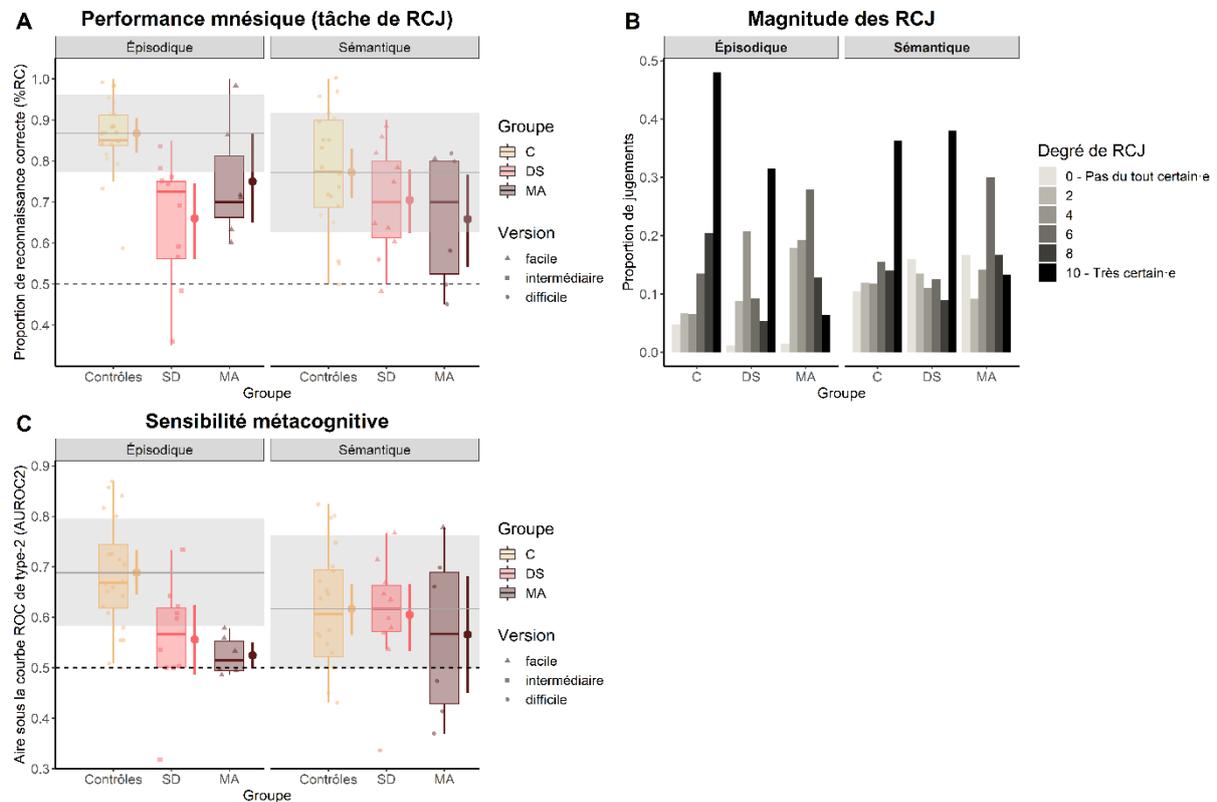
Dans la tâche épisodique, seuls les scores AUROC2 des contrôles ( $M = 0.69$ ,  $SD = 0.11$ ) étaient significativement différents de 0.5 ( $V = 210$ ,  $p < .001$ ) (pour les personnes DS :  $M = 0.56$ ,  $SD = 0.11$ ,  $V = 36$ ,  $p = 0.12$ ; pour les MA :  $M = 0.52$ ,  $SD = 0.04$ ,  $V = 15$ ,  $p = .44$ ). Sur la tâche sémantique, les scores AUROC2 des personnes DS ( $M = 0.60$ ,  $SD = 0.12$ ;  $V = 82$ ,  $p = .008$ ) et des contrôles ( $M = 0.62$ ,  $SD = 0.12$ ;  $V = 144$ ,  $p = .002$ ) étaient différents de 0.5, mais pas ceux des personnes MA ( $M = 0.57$ ,  $SD = 0.17$ ;  $V = 20$ ,  $p = .38$ ).

Pour la tâche épisodique, puisque les performances de premier ordre n'étaient pas équivalente entre les groupes, nous avons vérifié que cela ne pouvait pas être à l'origine des différences observées sur la sensibilité métacognitive. La corrélation entre le %RC et les AUROC2 n'était pas significative ( $r = 0.09$ ,  $p = 0.60$ ).

**Contrôle pour l'âge**

L'âge corrélait significativement uniquement avec les scores AUROC2 sur la tâche épisodique ( $r = -0.51$ ,  $p < .001$ ), mais pas sur les autres mesures ( $ps > .05$ ). Nous avons donc réalisé une ANCOVA avec l'âge comme covariable sur les scores AUROC2. En prenant en compte l'âge, l'effet du groupe sur les scores AUROC2 restait significatif ( $F(2,32) = 5.83$ ,  $p =$

.007,  $\eta_p^2 = 0.27$ ). Les scores AUROC2 des contrôles étaient significativement plus élevés que ceux des personnes DS ( $t(32) = 2.52$ ,  $p_{\text{bonferroni}} = 0.02$ ) et MA ( $t(32) = 2.91$ ,  $p_{\text{bonferroni}} = 0.006$ ). La différence des scores AUROC2 des personnes DS et MA n'était pas significative ( $t(32) = 1.15$ ,  $p_{\text{bonferroni}} = 0.26$ ).



**Figure 5.5 :** Résultats sur la tâche de RCJ.

A) Performance de premier ordre, B) Magnitude des RCJ, C) AUROC2. Les points épais sont les moyennes par condition, avec l'intervalle de confiance à 95%. Les points plus petits sont les données pour chaque participant. Les zones grises indiquent les scores situés à moins d'un écart-type de la moyenne des contrôles.

### 5.3.3. Questionnaire de plaintes mnésiques

Les scores des proches ont pu être obtenu pour 9 contrôles, 12 personnes DS et 4 personnes MA. Les scores de plaintes mnésiques différaient selon le groupe ( $F(2,37) = 5.55$ ,  $p = .008$ ,  $\eta_p^2 = 0.23$ ). Les personnes MA ont rapporté plus de difficultés mnésiques ( $M = 24.57$ ,  $SD = 4.08$ )<sup>5</sup> que les contrôles ( $M = 15.25$ ,  $SD = 6.47$ ) ( $t = 3.32$ ,  $p_{\text{bonferroni}} = .006$ ). Les scores des personnes DS ( $M = 18.23$ ,  $SD = 7.17$ ) n'étaient significativement pas différents de ceux des MA ( $t = 2.12$ ,  $p_{\text{bonferroni}} = .12$ ) ni des contrôles ( $t = 1.31$ ,  $p_{\text{bonferroni}} = 0.60$ ).

<sup>5</sup> Un score élevé indique une plus grosse plainte (i.e., plus de difficultés dans la vie de tous les jours)

La différence entre les scores des patients et ceux donnés par leurs proches étaient différentes selon les groupes ( $F(2,22) = 3.49, p = .048, \eta_p^2 = 0.24$ ). En effet, les personnes DS ont plus eu tendance à sous-estimer leurs difficultés mnésiques par rapport à ce que leurs proches rapportaient ( $M = -4.17, SD = 7.58$ )<sup>6</sup> alors que les contrôles ont eu tendance à plus surestimer leurs difficultés ( $M = 5.22, SD = 7.73$ ) ( $t(22) = 3.92, p_{\text{bonferroni}} = .002, d_z = 0.78$ ). La différence entre les scores des patients et ceux de leur proche ne différait significativement pas entre les DS et MA ( $M = -2.25, SD = 11.09$ ) ( $t(22) = 2.03, p_{\text{bonferroni}} = .16, d_z = 0.41$ ).

#### 5.4. Discussion du Chapitre 5

Le but de cette expérience était de mesurer les capacités métacognitives de personnes présentant un trouble de la mémoire sémantique, la DS, afin de voir la potentielle relation entre altération du système sémantique et les mesures de métacognition. Un second objectif était de voir s'il existait une double dissociation sur la sensibilité métacognitive, notamment des PoR, entre les personnes présentant plus de difficultés de mémoire épisodique (i.e., MA) et ceux présentant des symptômes plus sémantiques (i.e., DS). Nous faisons l'hypothèse que les participants DS auraient une moins bonne sensibilité métacognitive sur les tâches sémantiques qu'épisodiques, en particulier sur les tâches de PoR. Nous nous attendions à observer une double dissociation avec une plus faible sensibilité métacognitive sur les tâches sémantiques chez les personnes DS que MA, et un profil inverse sur les tâches épisodiques.

Nos résultats ne permettent pas de confirmer nos hypothèses. En effet, sur les tâches de PoR, la sensibilité métacognitive ne différait pas entre les groupes. De plus, sur les tâches de PoR et RCJ sémantiques, les scores AUROC2 significativement différents de 0.5 chez les personnes DS et les contrôles, montrent une bonne sensibilité métacognitive chez ces patients, même sur des tâches sémantiques. Ce résultat remet donc en question l'hypothèse selon laquelle un déficit d'une forme de mémoire (i.e., ici sémantique) entraînerait une moins bonne sensibilité métacognitive sur les tâches mettant en jeu cette forme de mémoire et même avec des jugements prédictifs. Ainsi, la relation observée entre déficit de type-1 et diminution de la sensibilité métacognitive observée dans des pathologies de la mémoire épisodique (Souchay, 2007) ne semble pas s'appliquer à cette pathologie de la mémoire sémantique. Une explication possible est que du fait de l'utilisation de versions différentes entre les groupes, les performances de premier ordre étaient en réalité plus élevées chez les DS que chez les contrôles et les MA sur la tâche de PoR sémantique, et cela n'était pas expliqué par les différences d'âge. Ainsi, bien que ces résultats ne permettent pas de mettre en évidence un déclin des performances métacognitives « pures » sur les tâches sémantiques, nous pouvons

<sup>6</sup> Ces scores sont la différence : score patient – score proche. Un score négatif suggère une sous-estimation des difficultés par les patients.

nous demander si les mêmes résultats auraient été obtenus avec une tâche plus compliquée, et donc des performances de premier ordre plus faibles, chez les personnes DS.

Ces résultats amènent à se questionner sur la relation entre déficit mnésique et sensibilité métacognitive. En effet, comme c'est le cas dans l'amnésie sans atteinte frontale (Janowsky et al., 1989; Shimamura & Squire, 1986), il ne semble pas qu'un déficit mnésique (même si pas observé sur la tâche de premier ordre ici du fait de l'utilisation de plusieurs versions) engendre systématiquement un déficit métacognitif sur les tâches mettant en jeu le type de mémoire affecté (i.e., sémantique dans le cas de la DS). Cela vaut également lorsque les participants doivent réaliser des prédictions au moment de la récupération en mémoire, qui font généralement plus intervenir les compétences mnésiques appropriées (e.g., recollection sur une tâche épisodique, récupération en mémoire sémantique ; Mazancieux et al., 2020). Malgré qu'ils aient reconnus moins de cibles que les contrôles sur la tâche de PoR épisodique, les patients DS et MA n'avaient pas une sensibilité métacognitive plus faible que celle des contrôles. Ce résultat chez les patients MA va à l'encontre des précédents travaux sur la MA (Ernst et al., 2015). Une explication possible est que la sensibilité métacognitive des participants contrôles n'était pas différente du hasard sur cette tâche, ce qui laisse supposer qu'elle était compliquée, même pour des personnes sans pathologie. Ainsi, la difficulté de cette tâche entraînait des performances métacognitives planchers puisque les jugements métacognitifs étaient aussi prédictifs de la performance de premier ordre que s'ils avaient été faits au hasard. Cet effet plancher peut expliquer donc pourquoi nous n'avons pas observé de différence entre les groupes sur cette tâche et limite donc les conclusions que nous pouvons tirer sur la relation entre déficit de premier ordre et sensibilité métacognitive.

Sur la tâche de RCJ épisodique, les personnes DS avaient une moins bonne sensibilité métacognitive que les contrôles, bien que cet effet ne soit que tendanciel après une correction statistique. Les performances de premier ordre n'étant pas significativement corrélées à la sensibilité métacognitive, ce résultat ne semblait pas expliqué par les différences de performances de premier ordre, comme cela aurait pu être suggéré par d'autres études montrant une relation entre la difficulté de la tâche de premier ordre et la sensibilité métacognitive (Galvin et al., 2003; Masson & Rotello, 2009; Vuorre & Metcalfe, 2021). Sur la base des modèles actuels, il semble donc difficile d'expliquer pourquoi les patients DS étaient potentiellement moins bons que les contrôles pour évaluer l'exactitude de leurs réponses épisodiques à posteriori. D'autres travaux seraient nécessaires pour approfondir cette question et vérifier s'il existe réellement un déficit de sensibilité du eRCJ dans la DS.

De plus, aucune dissociation de sensibilité métacognitive selon le type de déficit mnésique n'a été mise en évidence. En effet, les scores de sensibilité n'étaient pas significativement différents entre les personnes DS et les MA, que ce soit avec des PoR ou des RCJ, que les tâches soient épisodiques ou en sémantiques. Cela peut s'expliquer par le fait que les

personnes DS avaient des performances de premier ordre plus faibles que les contrôles sur les tâches épisodiques et non statistiquement différentes des personnes MA. Cela montre que les personnes DS avaient des difficultés sur les tâches épisodiques également.

Les résultats obtenus sur les tâches de métacognition fournissent une mesure objective de l'*awareness* que les patients ont de leurs capacités mnésiques. Ce type de mesure est complémentaire à des évaluations plus subjectives de l'*awareness* (e.g., Cosentino et al., 2011; DeLozier & Davalos, 2016). Nous avons donc complété les observations faites sur les tâches de métacognition par un questionnaire de plaintes mnésiques. L'utilisation de la méthode des scores de décalage entre patients et proches sur des questionnaires a permis de mettre en évidence que les personnes DS sous-estimaient leurs difficultés dans la vie de tous les jours. Ce résultat contraste avec l'observation que les personnes DS reçus en consultation se plaignent généralement de leur difficulté à identifier des personnes, à reconnaître et à nommer certains objets, ainsi qu'à comprendre des mots isolés (Belliard et al., 2007), qui laisse supposer que les personnes DS ne sont pas anosognosiques. La majorité des questions du questionnaire utilisé (McNair & Kahn, 1983) portent sur la mémoire épisodique (e.g., « Vous arrive-t-il de répéter plusieurs fois la même chose parce que vous oubliez l'avoir déjà dite ? ») ou de désorientation (e.g., « Vous arrive-t-il de vous perdre dans des lieux familiers ? »), et seulement deux sur 20 portent sur la mémoire sémantique (e.g., « Oubliez-vous le nom ou le prénom des personnes qui vous sont familières ? »). Les sujets âgés sains semblent au contraire surestimer leurs difficultés. Les personnes DS sous-estiment leurs difficultés épisodiques bien que celles-ci ne soient pas nécessairement déficitaires. Une explication possible est que les patients DS ont effectivement une bonne conscience de leurs déficits sémantiques (e.g., difficulté à retrouver des noms) mais ils prennent moins conscience de leurs comportements mnésiques lorsque ceux-ci sont en lien avec des difficultés plus au second plan dans leur pathologie (e.g., épisodiques).

Pour conclure, les modèles d'anosognosie actuels (e.g., CAM, Agnew & Morris, 1998; Hannesdottir & Morris, 2007; Morris & Mograbi, 2013) mettent en avant l'importance de la mémoire épisodique pour garder une bonne conscience de soi et une métacognition optimale (voir Figure 5.1). De plus, la mémoire sémantique interviendrait dans une forme autobiographique car la mise à jour régulière de la base de données personnelles jouerait également un rôle important. Les personnes DS n'ont généralement pas de difficultés à se souvenir des épisodes vécus, en particulier s'ils sont récents (Hodges & Patterson, 2007; Hou et al., 2005; Ivanoiu et al., 2006; Piolino et al., 2003). Les connaissances qu'ils ont sur eux-mêmes (i.e., mémoire autobiographique sémantique) sont également plutôt préservée lorsqu'elles concernent leur soi du présent (Duval et al., 2012). Sur la base du modèle CAM, cela pourrait expliquer pourquoi les patients DS ne présentent pas de déficit métacognitif

général. Une limite des modèles de l'anosognosie est qu'ils restent très généraux et ne prennent pas en compte les différentes façons d'évaluer sa mémoire (e.g., jugements prédictifs ou rétrospectifs). Ainsi, les modèles de métacognition permettent une description plus fine des processus mis en jeu dans le monitoring (Koriat, 1993, 1997). En se basant sur ces modèles pour évaluer des aspects plus précis de l'*awareness*, notre étude ne permet pas de mettre en évidence un déclin de la sensibilité métacognitive général, ni spécifique aux tâches sémantiques, et amène à se questionner sur la nécessité du système sémantique dans le monitoring.

## CHAPITRE 5 : CONCLUSION

Nos résultats ne permettent pas de démontrer un déclin de la sensibilité métacognitive qui soit relié à un déficit sémantique puisque les personnes DS arrivaient aussi bien que les contrôles à prédire leurs performances de reconnaissance, que ce soit sur une tâche sémantique ou épisodique. En revanche, lorsqu'ils devaient évaluer rétrospectivement l'exactitude de leurs réponses, les patients DS étaient moins sensibles que les contrôles, seulement lorsque la mémoire épisodique était testée. Ainsi, les patients DS ne présentent pas de déficit métacognitif général ni spécifique aux tâches mettant en jeu la mémoire sémantique.

De plus, en comparant des personnes DS à des MA, nous n'avons pas mis en avant de double dissociation sur la sensibilité métacognitive, que ça soit avec des jugements PoR ou des RCJ.

Mis ensemble, ces résultats remettent en question le **lien entre déficit de premier ordre et sensibilité métacognitive**, qui est notamment suggéré par les études sur la maladie d'Alzheimer.

# Chapitre 6. Importance du traitement sémantique pour les prédictions en mémoire épisodique

---

---

Les recherches sur la métacognition montrent que les prédictions d'une reconnaissance future (PoR) sont aidées par le traitement sémantique et reposent sur l'accès aux propriétés sémantiques de l'information qui doit être mémorisée (voir [Chapitre 3.2.3](#)). Dans le chapitre précédent, nous n'avons pas pu mettre en évidence qu'un trouble sémantique soit à l'origine d'un déficit métacognitif particulier. Dans ce chapitre, nous avons voulu évaluer si, chez les personnes saines, le traitement sémantique de paires de mots est nécessaire pour prédire avec précision une récupération future en mémoire. Dans 4 expériences, les participants ont effectué une tâche de mémoire épisodique dans laquelle leur mémoire de paires indice-cible a été testée. Ils devaient prédire leur capacité à reconnaître la cible en faisant une PoR sur présentation de l'indice. L'accès au traitement sémantique des indices et des cibles a été manipulé par la présentation de deux types d'items : sans signification (i.e., pseudo-mots) ou ayant du sens (i.e., vrais mots). Dans l'ensemble, les résultats montrent que les PoR sont influencées par le type d'indice puisque les participants ont prédit reconnaître moins de cibles lorsqu'ils voyaient un indice sans signification. Lorsqu'ils voyaient des mots réels comme indices, ils prédisaient reconnaître plus de cibles ayant du sens. Ce résultat était aussi observé dans des conditions où les cibles sans signification étaient en fait mieux reconnues que les cibles ayant du sens. De plus, le traitement sémantique de la cible et de son indice était important pour maximiser la sensibilité métacognitive. Cependant, en utilisant un paradigme standard de FOK dans l'expérience 6.4, nous avons montré que l'effet du traitement sémantique sur la sensibilité métacognitive résultait du rappel plus élevé de paires d'items ayant du sens (i.e., paires mot - mot). Lorsque les cibles n'étaient pas rappelées, aucun effet du traitement sémantique sur la sensibilité métacognitive n'était observé. Ces résultats soutiennent l'idée que le traitement sémantique des informations est important pour prédire une récupération future en mémoire épisodique, et principalement car il aide à se rappeler des cibles, et qui est un indice utilisé pour les PoR.

Ce chapitre fait l'objet d'un article en cours d'écriture, en anglais. Il a été reformaté pour les besoins de ce manuscrit.

**Devaluez, M.**, Moulin, C. & Souchay, C. (in prep). Predicting recognition memory of meaningless words: the importance of semantic processing for episodic metacognition.

Les résultats de l'Expérience 6.1 ont été présentés au 1<sup>er</sup> colloque du GDR Mémoire, Cahors, France, Octobre 2021.

## 6.1. Introduction

Even though memory retrieval is not always successful, people are often still able to successfully predict subsequent retrieval outcomes, an ability which refers to metacognition (Flavell, 1979). In studies of metacognition, participants either predict their upcoming performance or judge the accuracy of their memories retrospectively. In studies of metacognition in memory, two paradigms dominate the literature, prediction of future recall (e.g., *judgment-of-learning*, JOL; Arbuckle & Cuddy, 1969) or future recognition (e.g., *feeling-of-knowing*, FOK; Hart, 1965). Such predictions about future memory retrieval are based on a variety of cues (Koriat, 1997) such as semantic knowledge (e.g., knowledge about objects' properties; Thomas et al., 2012). Previous work has shown that the semantic attributes of stimuli influence metacognitive predictions of episodic memory performance. However, meaning and semantic attributes of to-be-remembered material sometimes less apparent. Such occasions, must occur when encountering novel material in language learning or unfamiliar contexts (e.g., proper nouns in new locations). Moreover, in neurological diseases in which semantic information is altered (e.g., semantic dementia; Snowden et al., 1989, 2004) a question arises as to whether episodic metacognition could also be influenced by semantic processing alteration. As such, it is possible that semantic memory deficits could also alter people's ability to predict their episodic memory. In the present work, we therefore aimed at better understanding the importance of one type of information on one's ability to predict future episodic memory performance: the access to semantic information about the verbal material to be remembered.

Episodic memory refers to the subjective remembering of personally experienced events (Tulving, 1984, 1985a, 1993). It has frequently been studied in comparison to semantic memory (Baddeley, 2001; Tulving, 1972), which relates to the declarative knowledge about the world and stores information about facts, objects and concepts independently from their original learning context. The two types of memory are interactive and episodic memory can be influenced by semantic processing (Renoult et al., 2019; Renoult & Rugg, 2020; Tulving, 1983). Semantic memory is formed from the extraction of statistical regularities from different experiences, and is reactivated when experiencing the world. For example, reading a word

does not only rely on orthographic and phonological decoding but also on a lexical-semantic system which activates semantic knowledge (e.g., Harm & Seidenberg, 2004; Pattamadilok et al., 2017; Seidenberg & McClelland, 1989). Such access to semantic information can be manipulated by lexicality, word frequency, and concreteness. It was proposed that recollection (i.e., the active process of remembering past specific details), which is at the core of episodic memory, is the reinstatement of the conceptual processing that happened during the experience in addition to the contextual information specific to the episode (see Renoult et al., 2019). Semantic processing consists for example in the access to categories, sensory information (e.g., visual, tactile, motor or auditory), and relationships between concepts.

Multiple examples of influence of semantic processing on episodic memory can be highlighted (see [Chapter 3.2.2](#) for details). For example, if intact, prior semantic knowledge facilitates episodic encoding and retrieval in amnesic patients (Kan et al., 2009). Other work shows that word pairs are better recalled when semantically related (e.g., shoe - boot) than unrelated (e.g., fence - helicopter)(Epstein et al., 1975). Moreover, according to the level-of-processing theory, memory is better for stimuli processed in relation to their semantic properties at encoding (i.e., deep encoding) compared to shallowly processed stimuli as it allows new information to be linked more easily to existing knowledge ( Craik & Lockhart, 1972; Lockhart & Craik, 1990). However, there is yet to evaluate if semantic processing is necessary not only for episodic memory (i.e., first-order performance) but also for one's ability to reflect on it (i.e., second-order performance).

Metacognitive judgments made on episodic tasks can be influenced by semantic memory in different ways (see [Chapter 3.2.3](#) for details). When asked to predict later cued-recall (JOL paradigm), participants gave higher JOLs for semantically related pairs compared to unrelated pairs (Castel et al., 2007; Connor et al., 1997; Dunlosky & Matvey, 2001; Mueller et al., 2013; Undorf & Erdfelder, 2015) even in cases in which actual memory performance was better for unrelated pairs (Carroll et al., 1997). Studies have also shown that participants are better at predicting their future recall abilities for related pairs than for unrelated pairs (Hertzog et al., 2002).

In FOK tasks, participants make predictions of future recognition after a failed recall attempt (Hart, 1965). FOKs are influenced by two heuristics: the subjective familiarity with the cue (Metcalf et al., 1993; Reder, 1988; Reder & Ritter, 1992; Schwartz & Metcalf, 1992) and the partial retrieval of information about the target that is to be retrieved (Koriat, 1993, 1994, 1995). According to an interactive model of both accounts (Koriat & Levy-Sadot, 2001), predictions are first influenced by the cue familiarity and if familiarity is sufficient, people can engage in research of target information which can accumulate to influence predictions subsequently. Studies have shown that even when the target does not come to mind, people can still use some partial information that they can retrieve about it such as semantic

characteristics (Koriat et al., 2003; Thomas et al., 2012). For example, Thomas et al. (2012) showed that the quality of accessible semantic features about the target influenced FOK. Indeed, participants gave higher and more accurate FOKs when they provided information about the affective connotation (i.e., positive or negative) of the target compared to when they could not give that information, and even more so if that information was correct. In another study (Experiment 2 of Hosey et al., 2009), participants learnt new face-name associations as well as fictional facts about the people (e.g., occupation, pet) and later made FOK for the names when presented with the faces. They predicted recognizing more names when they were able to recall semantic attributes learnt about the people (e.g., “she is a doctor”). The same results were found when using drawings of imaginary animals associated with names and specific information to learn (Schwartz et al., 2014). These results show once more that semantic features can be used to guide predictions. However, in these experiments features were not semantic per se but rather semantic-like or even episodic as they were learnt in the encoding phase. Moreover, FOK are also affected by levels of processing (Lupker et al., 1991; Thomas et al., 2012). In their experiment (Experiment 1 of Thomas et al., 2012), pairs that were deeply encoded (i.e., in relation to their meaning, semantically) were given higher and more accurate FOK than shallowly encoded (in relation to perceptual features) pairs. But what happens when such semantic information is not available?

Meaningless items such as pseudowords logically do not have any representation stored in memory before task completion. In addition to being meaningless, pseudowords are new stimuli that participants have never encountered in their life before completing the task and are by nature less familiar than real words. As cue-familiarity has been shown to be correlated with predictions' accuracy (Metcalfe et al., 1993), we wondered how using meaningless words as cues could affect predictions. Meaningless materials have already been used in studies on predictions of recognition. For example, in two studies (Koriat, 1993) participants learnt strings of letters. After a few seconds, participants recalled as many letters as possible and made an FOK (i.e., predicted the likelihood that they would be able to identify the correct target). Each FOK was followed by an 8-alternative forced-choice recognition test. Results revealed significant positive correlations between FOK and actual performance (i.e., the FOK reflects actual memory performance above chance). In a third experiment of the same paper (Koriat, 1993), participants performed the same kind of task but learned pairs of nonsense cues with meaningful targets but no measure of metacognitive sensitivity was provided. The main goal of these three experiments was not to assess the importance of semantics on FOK but rather explore the role of accessibility to features about the target in predictions. Moreover, one study (Experiment 3 of Thomas et al., 2012) explored FOK for meaningless shapes as a means to evaluate the influence of access to perceptual features on FOK when there is reduced access to conceptual features. However, no comparison between meaningless and meaningful items

was provided in any of these studies. Therefore, in the present work we wanted to evaluate the specific implication of semantics of the cue and of the target separately on metacognitive measures (i.e., magnitude, sensitivity) of predictions by using pairs of meaningless and meaningful items.

Multiple measures of metacognition can be investigated as they are complementary and provide different information (Schraw, 2009, see [Chapter 1.2.2](#)). In the experiments reported here, magnitude of judgments and metacognitive sensitivity will be taken into account, as preregistered, as they tap different metacognitive abilities which could be distinctively influenced by semantic processing.

To the best of our knowledge, no study has yet directly compared the accuracy of predictions of long-term memory for meaningful and meaningless materials, nor if the role of meaning was specific to the cue or the target of a learnt pair. The goal of this project was to assess whether semantic processing (i.e., properties related to meaning) of the cue and the target of a pair is necessary to accurately predict future episodic recognition memory. In four experiments (see Tableau 6.1), we studied the implications of limiting semantic processing of the material on predictions of recognition memory (PoR) by using pseudowords and real words. Our rationale is that real words are meaningful materials which allow semantic processing and pseudowords are meaningless materials without semantic properties.

In the first 3 experiments, we assessed participants' ability to predict their performance without explicitly asking them to access the target when making their PoR. Recall might be harder for meaningless items and previous work has shown that asking participants to retrieve the target before predicting their performance leads to better metacognitive sensitivity (Kelemen et al., 2000). Therefore, here, the goal was to measure people's ability to predict their performance on a recognition test without necessarily prompting them to recall in order to guide their predictions. To do so, we asked participants to first learn all the cue-target pairs. Then, they predicted future memory for the targets when prompted with the cues in a second phase (as in a delayed-JOL paradigm, e.g., Dunlosky & Nelson, 1992; Nelson & Dunlosky, 1991) with the difference that our final task was recognition. Memory was tested directly after their PoR for each trial on a two-alternative forced-choice test. As we suspected recall abilities to still play a role in making PoR even when participants were not explicitly asked to recall the targets, in a fourth experiment, we added a recall attempt phase before predictions were made to evaluate the role of recall abilities in such predictions, therefore using a classical FOK task.

We hypothesized that semantic processing provides a framework for learning but also for the use of cues influencing PoR. When there is no access to meaning of the target, one can only rely on other types of cues (e.g., orthographic or phonological cues) to make predictions which might not be as efficient. Moreover, according to the cue-familiarity hypothesis,

predictions should also be influenced by knowledge of the cue. As such, we expected metacognition to be affected by the lack of possible access to semantic information of both the cue and the target. In Experiment 6.1, we aimed at evaluating people's ability to predict their recognition memory of targets with and without semantic properties which were cued by meaningful or meaningless cues. Meaningfulness affected both metacognitive bias and sensitivity. Unexpectedly, in Experiment 6.1, we found that recognition performance was worse for meaningless targets and for targets cued by meaningless cues. However, second-order performance is related to first-order performance (Fleming & Lau, 2014; Vuorre & Metcalfe, 2021). Therefore, in two other experiments (Experiments 6.2 and 6.3), we aimed at investigating whether the influence of semantic processing on metacognition came from processing differences at play at encoding (Experiment 6.2) and from familiarity discrepancies (Experiment 6.3). We wanted to investigate if the same results would be obtained when helping processing of meaningless items at encoding without using semantics, as quality of original encoding has been shown to influence PoR (Experiment 1 of Thomas et al., 2012). Therefore, we wondered if the effect of meaninglessness on metacognitive measures was a consequence of poorer encoding and/or lower first-order performance or if even when helping processing of such material in other ways (i.e., increased time for learning), semantic processing is still important for accurate PoR. In Experiment 6.2, we increased the presentation time of meaningless items during the encoding phase. By doing so, we attempted to reduce the discrepancy in first-order performance between words and pseudowords and favoring the encoding of pseudowords (see Thomas et al., 2011 for an example of study in which this manipulation was used to decrease the difference between groups performance). We explored whether allowing more time for the processing of pseudowords cues and targets reduced the effects observed in Experiment 6.1 on the PoR and on metacognitive sensitivity. In Experiment 6.3, we increased familiarity with meaningless items by making participants read the pseudowords before paired-associates learning. The goal was to check if by increasing familiarity of pseudowords (via prior exposure) the effects of meaninglessness on metacognition disappeared. Finally, we expected the meaninglessness effect on metacognitive bias and sensitivity to come, at least in part, from the difference in memory strength and recall between meaningful and meaningless items (i.e., memory constraint hypothesis; Hertzog et al., 2010). Therefore, in Experiment 6.4 we wanted to explore whether effects observed in the first 3 experiments were a result (at least in parts) of recall abilities. Based on previous research (Schwartz et al., 2016), we expected less effect of lack of semantic processing on PoR when only considering trials in which participants failed to recall the target.

The hypotheses and the methods of the 4 experiments of this project were pre-registered on the Open Science Framework (<https://osf.io/3zaqy/>). For the sake of clarity, we decided not

to include results of the bias measure here as there are redundant with results from first-order performance and the magnitude of the judgments.

**Tableau 6.1** : Summary of the 4 experiments of Chapter 6.

Experiment	Manipulation	Goal
Experiment 6.1	No manipulation	
Experiment 6.2	Presentation time: longer time for pseudowords	Help processing of pseudowords at encoding
Experiment 6.3	Familiarization phase: presentation of the pseudowords before encoding	Increase familiarity with the pseudowords
Experiment 6.4	Addition of a recall attempt before making the PoR	Assess the role of recall

## 6.2. Experiment 6.1: Predicting recognition of meaningless words

This first study aimed at evaluating if the ability to predict performance on an episodic recognition test changes when semantic information is manipulated. Previous research conducted with the old/new recognition paradigm has shown no difference in memory performance between words and pseudowords (Gardiner & Java, 1990; Rajaram et al., 2002 ; Experiment 1 of Geraci et al., 2009; but see Experiment 2 of Geraci et al., 2009 in which pseudowords were less recognized than words), although responses might rely on distinct mechanisms (but see Haaf et al., 2021 for contradictory results regarding underlying processes). Therefore, we did not expect any effect of cue and target types on recognition memory. However, because predictions can be influenced by cue-familiarity (Metcalf et al., 1993; Reder, 1988; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992; Son & Metcalfe, 2005) which might be related to cue frequency (Noble, 1954; Tanaka-Ishii & Terada, 2011), we expected lower PoR for targets cued by pseudowords than for targets cued by words. Moreover, we hypothesized that semantic processing of both the target and the cue of the pair is important for accurate metacognition as it provides more information to rely on when making predictions. It also provides a better opportunity for deep encoding as it is possible to use semantic information which could in turn influence predictions' accuracy (Thomas et al., 2012). As such, we expected greater metacognitive sensitivity and efficiency for pairs with word targets as well as for pairs with word cues.

### 6.2.1. Method

#### **Participants**

We recruited 76 participants for this experiment ( $M_{age} = 19.99$ ,  $SD_{age} = 1.75$ , 8 males). All participants were French-speaking Psychology students from the University of Grenoble-Alpes who received course credit in exchange. None reported any history of neurodevelopmental, neurological or psychiatric disorder. The required sample size was determined with G\*power software for a medium effect size  $d_z = 0.5$  and a statistical power of .99. Data from 10 participants were excluded from the analyses (see [Results](#) section for details). Therefore, data from 66 participants were included in the analyses ( $M_{age} = 19.95$ ,  $SD_{age} = 1.55$ , 7 males). All participants provided informed consent before study completion.

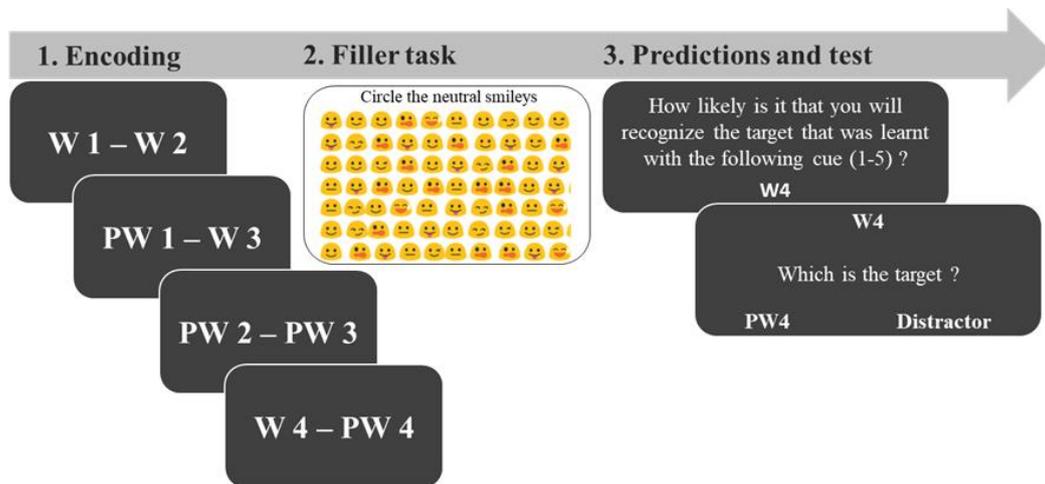
#### **Material**

We selected 180 two-syllable concrete French words from the French Lexique database (New et al., 2004). We created 180 two-syllable pseudowords (i.e., strings of letters which could be read but are meaningless in French; e.g., *cituse*). To limit any visual or phonological effects, the mean number of phonemes, the mean number of letters, and the mean frequency of bigrams of the list of pseudowords were matched to the list of words. Cues and/or targets were meaningful real words or meaningless pseudowords. Participants studied four types of pairs: two words (WW; e.g., *perruque-rectangle*), two pseudowords (PW/PW; e.g., *atue-rattie*), a pseudoword cue with a word target (PW/W; e.g., *cituse-drapeau*), and a word cue with a pseudoword target (W/PW; e.g., *grenier-misanclé*), see [Appendix D](#) for more examples. The task was implemented on E-Prime 3 software (Psychology Software Tools, 2016).

#### **Procedure**

Participants performed two blocks of the same task (see Figure 6.1). Each block consisted of 3 parts. First, participants learnt 60 cue-target pairs which were each presented for 2s and followed by a 500ms inter-stimulus interval. After learning, participants completed a visual exploration task (i.e., filler task). Finally, the cues were presented in a random order and for each, participants had to predict how likely they were to recognize the correct target between two alternatives, the PoR. This was made on a five-point Likert scale (1 = very unlikely; 5 = very likely). After each judgement, participants selected the target in a two-alternative forced-choice test.

Participants were tested on 30 pairs of each type (15 in each block) presented in a random order. In the recognition test, distractors were from the same type as the target (i.e., the distractor was a pseudoword if the target was a pseudoword). Words in word-word pairs were semantically unrelated. Cues and targets were counterbalanced across participants.



**Figure 6.1 :** Procedure of one block of Experiment 6.1.

Participants performed two blocks with different items. Pairs were presented in a random order.  
W = word; PW = pseudoword

### Analyses

Analyses were performed on the overall performance for the 120 trials (across the two blocks). All analyses were performed on Rstudio (version 1.3.1073) using 2 (cue-type: word, pseudoword) \* 2 (target-type: word, pseudoword) repeated measures ANOVAs as pre-registered. For completeness, we also analyzed each effect with linear models using the `lm()` function from *stats* package. Measures of interest were first-order performance, magnitude, and metacognitive sensitivity (gamma scores and AUROC2). These will be the same for the other three experiments. All analyses were conducted using  $\alpha = .05$ . If not indicated otherwise, analyses met normality prerequisites for the use of linear models.

Memory (i.e., first order) performance was measured with the proportion of correct recognition (%CR) as well as the signal detection theory (SDT) discrimination index  $d'$  (Green & Swets, 1966). To use a similar scale range as performance (i.e., 0 to 1) and because the recognition test was in a 2-alternative forced choice (2AFC) format, the magnitude of PoR was analyzed by transforming predictions of future recognition ranging from 1 (very unlikely) to 5 (very likely) to range from 0.5 to 1. As such, a prediction of 1 was attributed a value of 0.5, 2 a value of 0.625, 3 a value of 0.75, 4 a value of 0.875, and 5 a value of 1. Two non-parametric measures were used to evaluate metacognitive sensitivity: standard gamma correlations (Goodman & Kruskal, 1954) and the type-2 area under ROC (receiver operating characteristic) curve (AUROC2, not preregistered). AUROC2 relies on SDT and is more bias-free than gamma correlations as it is less influenced by the tendency to give high or low predictions (see Fleming & Lau, 2014). AUROC2 of 0.5 indicates at-chance metacognitive performance. We also attempted to evaluate metacognitive efficiency, which is more bias-free and less performance-dependent than gamma correlations and AUROC2. We used the  $M_{ratio}$  index from the meta- $d'$  framework. Based on signal detection theory (SDT),  $M_{ratio}$  provides an estimation of

metacognitive efficiency by computing the ratio between first-order performance ( $d'$ ) and metacognitive performance (meta- $d'$ ). Meta- $d'$  was determined using the HMeta-d toolbox (<https://github.com/metacoglab/HMeta-d>) in R which is a Bayesian framework for estimating meta- $d'$  (Fleming, 2017). We tried to estimate the effects of cue-type and target-type on metacognitive sensitivity with the “Function\_metad\_2wayANOVA” function. However, these data were not conclusive (as explained below in the Results section) and could not be interpreted. As we preregistered analyses on Mratio, we still provide outputs of the models in [Appendix I](#).

## 6.2.2. Results

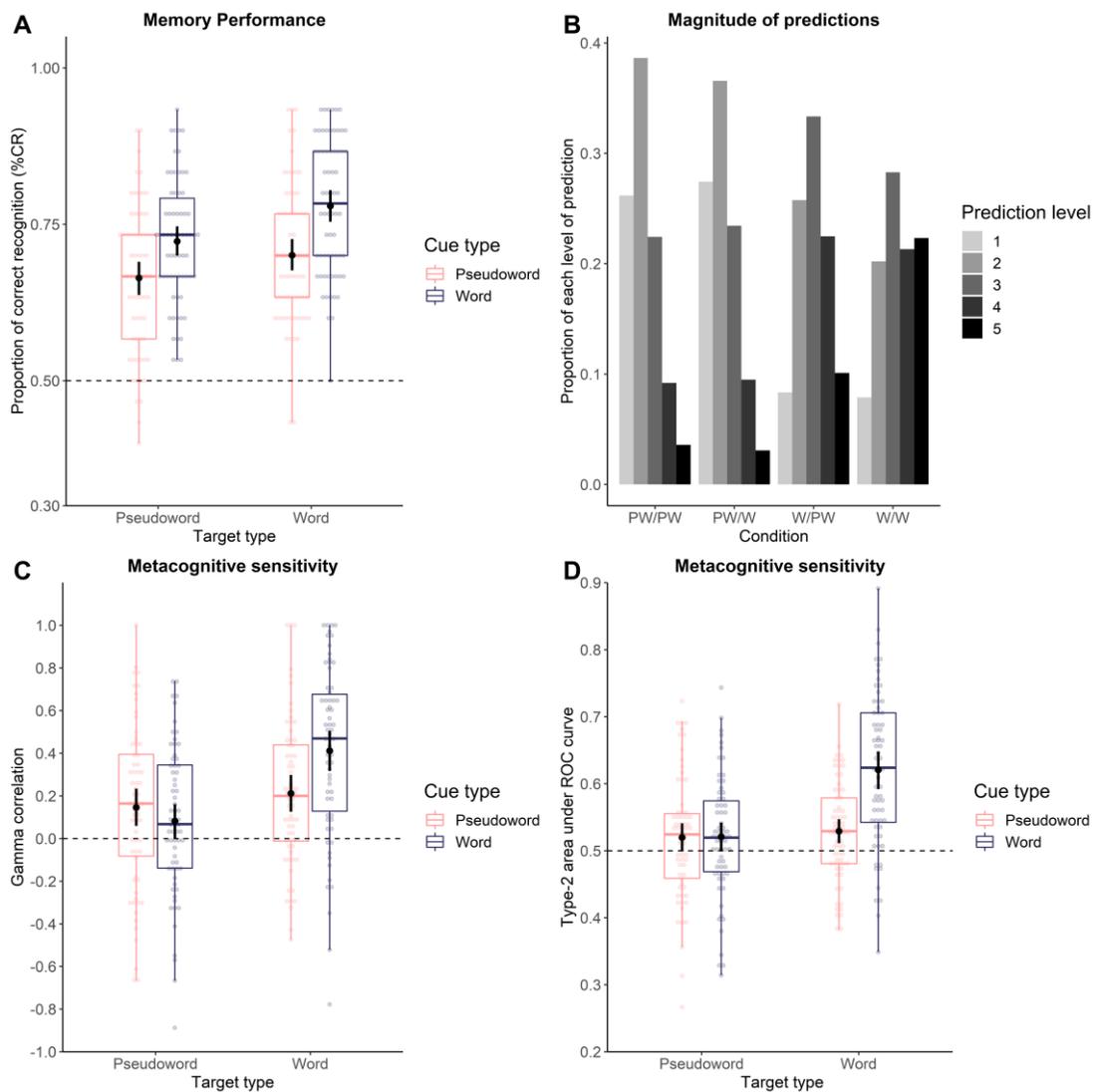
### **Data exclusion**

The %CR in each block for each participant was used as a criterion for inclusion in the analyses. We preregistered that we would exclude from the analyses participants who perform below 55%CR in at least one of the four experimental conditions. However, we noticed that using this threshold in each condition would create a bias in the results: 14 participants performed below 55%CR in the condition in which both the cue and the target were pseudowords but fewer participants performed below 55%CR in the other conditions (4 in the PW/W condition and W/W condition each; 5 in the W/PW condition). Therefore, removing participants on this criterion meant excluding participants who were most sensitive to the effect we are studying (i.e., worst performance with meaningless items). Instead, we checked whether participants performed overall above chance during the completion of the block than if they performed well in each condition which could just reveal a strategy used. We excluded 5 participants performing below 55%RC in at least one block. Five more participants were excluded as they performed close to ceiling (superior to 95%CR) in the W-W condition, which prevented the calculation of metacognitive sensitivity scores. To provide information about first-order performance for participants that were included in analyses on second-order measures (e.g., gamma scores), we excluded participants with ceiling effects from analyses of all measures. Note that removing such participants might lead to an underestimation of the actual effect of semantic processing on recognition performance.

### **Memory performance**

Participants recognized more word than pseudoword targets ( $F(1,65) = 15.54, p < .001, \eta_p^2 = 0.19 [0.05, 0.36]$ ), see Figure 6.2a. The type of cue also had an influence on recognition memory ( $F(1,65) = 39.32, p < .001, \eta_p^2 = 0.38 [0.20, 0.53]$ ) as targets cued by a word in the encoding phase were better recognized than targets cued by a pseudoword. The effect of target-type did not significantly interact with cue-type ( $F(1,65) = 0.96, p = .33, \eta_p^2 = 0.01 [0,$

0.12]). We observed the same pattern of results using  $d'$  as a measure of first-order performance. See [Appendix E](#) for details of means and standard deviations.



**Figure 6.2 : Results of Experiment 6.1.**

A) First-order performance, B) Magnitude of judgments, C) Gamma scores, and D) AUROC2 scores.

Each dot represents one participant. Black dots and black lines are the means and the 95% confidence interval around it. For metacognitive bias, scores below zero suggest underestimation of performance and scores above zero suggest overestimation.

### ***Magnitude of judgments***

Magnitude was analyzed using the transformed predictions (i.e., 0.5 to 1). PoR were higher for word targets than for pseudoword targets ( $F(1,65) = 32.85$ ,  $p < .001$ ,  $\eta_p^2 = 0.34$  [0.16, 0.49]). Participants were also affected by the nature of the cue when making their PoR ( $F(1,65) = 149.64$ ,  $p < .001$ ,  $\eta_p^2 = 0.70$  [0.57, 0.78]): PoR were higher when participants saw a word cue than a pseudoword cue (see [Appendix E](#)). These results are in keeping with the memory performance above. However, there was also a significant interaction effect ( $F(1,65) = 49.37$ ,

$p < .001$ ,  $\eta_p^2 = 0.43$  [0.25, 0.57]) as PoR depended on the nature of the target only when the cue was a word ( $t(65) = 7.34$ ,  $p < .001$ ,  $d_z = 0.90$  [0.62, 1.20]) but not when it was a pseudoword ( $t(65) = 0.49$ ,  $p = .63$ ,  $d_z = 0.06$  [-0.18, 0.30]<sup>7</sup>). When the cue was a word, participants made higher PoR for word targets than for pseudoword targets. Proportions of each raw judgment, across the 30 trials of each condition, are provided in Figure 6.2b.

### **Metacognitive sensitivity and efficiency**

**Gamma correlations.** Gamma correlations were higher for word targets than for pseudoword targets ( $F(1,65) = 21.65$ ,  $p < .001$ ,  $\eta_p^2 = 0.25$  [0.09, 0.41]). Gamma scores were not significantly influenced by the nature of the cue ( $F(1,65) = 1.77$ ,  $p = .19$ ,  $\eta_p^2 = 0.03$  [0, 0.14]). However, the effect of target-type interacted with cue-type ( $F(1,65) = 10.96$ ,  $p = .002$ ,  $\eta_p^2 = 0.14$  [0.02, 0.30], see Figure 6.2c). Post-hoc analyses showed that gamma correlations were lower for pseudoword targets only when the cues were words ( $t(65) = 5.98$ ,  $p < .001$ ,  $d_z = 0.74$  [0.47, 1.01]) but were not affected by the nature of the target when cues were pseudowords ( $t(65) = 1.07$ ,  $p = 0.29$ ,  $d_z = 0.13$  [-0.11, 0.38]). For pairs with word cues, gamma scores were higher when the target was also a word than when it was a pseudoword. Complementary analyses (not pre-registered) were performed using one sample t-tests which revealed that gamma scores were different from zero in all conditions (PW/PW:  $t(65) = 3.12$ ,  $p = .003$ ,  $d = 0.36$ ; PW/W:  $t(65) = 5.06$ ,  $p < .001$ ,  $d = 0.62$ ; W/W:  $t(65) = 8.35$ ,  $p < .001$ ,  $d = 1.03$ ), except in the W/PW condition ( $t(65) = 1.93$ ,  $p = 0.058$ ,  $d = 0.24$ ).

**Type-2 AUROC curves.** Replicating the gamma analysis, AUROC2 were higher for word targets than pseudoword targets ( $F(1,65) = 24.60$ ,  $p < .001$ ,  $\eta_p^2 = 0.27$  [0.11, 0.44], see Figure 6.2d). Contrary to gamma scores, AUROC2 were influenced by the nature of the cue ( $F(1,65) = 13.52$ ,  $p < .001$ ,  $\eta_p^2 = 0.17$  [0.04, 0.34]). AUROC2 were higher for pairs in which targets were cued by a word than for pairs with a pseudoword cue. The interaction between the nature of the target and its cue was again significant ( $F(1,65) = 19.81$ ,  $p < .001$ ,  $\eta_p^2 = 0.23$  [0.08, 0.40]). Post-hoc analyses revealed that the nature of the target only had an impact on AUROC2 when the cues were words ( $t(65) = 6.90$ ,  $p < .001$ ,  $d_z = 0.85$  [0.57, 1.14]) but not when the cues were pseudowords ( $t(65) = 0.60$ ,  $p = .55$ ,  $d_z = 0.07$  [-0.17, 0.32]), in line with the gamma analysis. AUROC2 for pairs with word targets were higher when the cue was also a word ( $M = 0.62$ ,  $SD = 0.11$ ) than when the cue was a pseudoword ( $M = 0.53$ ,  $SD = 0.57$ ).

**Metacognitive efficiency.** As pre-registered, we attempted to evaluate the effects of target-type, cue-type, and their interaction on metacognitive efficiency by estimating these effects on group

<sup>7</sup> The residuals of the model were not normally distributed. Using non-parametric Wilcoxon test:  $V = 756$ ,  $p = .44$ .

Mratio (meta- $d'/d'$ ) using hierarchical Bayesian framework (Fleming, 2017; “Function\_metad\_2wayANOVA” function). However, none of the 3 models converged thus preventing interpretation. Therefore, we did not report these results here but graphs of the output of the models are in [Appendix I](#).

### 6.2.3. Interim discussion

In this task, memory and metamemory variables were both influenced by the type of cues and targets. The magnitude of predictions did not dissociate however from the effects on memory: roughly parallel effects on memory performance and memory predictions were found. In addition, this first experiment shows that metacognition of memory for meaningless material is possible (gamma correlations are reliably above zero). However, metacognitive sensitivity does vary according to the type of material studied: participants were less metacognitively sensitive when at least one item of the pair was meaningless. Taken together, this suggests that PoRs were influenced by the type of cue, and when the cue was meaningful participants also used the type of target to infer their PoR. This result will be discussed in the General Discussion.

Surprisingly, recognition memory of targets was better when they were associated with meaningful cues than with meaningless cues. One plausible explanation is that pseudowords were read slower than known words (Valdois et al., 2006; Weekes, 1997) during the learning phase. Therefore, because targets and cues were presented together and all pairs were presented for the same amount of time (i.e., 2s), targets following a pseudoword might not have been processed for a sufficient amount of time.

Because first-order performance on a task (i.e., recognition memory here) correlates with and influences metacognitive sensitivity, especially when guessing is allowed (Fleming & Lau, 2014; Vuorre & Metcalfe, 2021), these effects in metamemory sensitivity need interpreting in the context of performance differences between the different types of pairs. As such, we wondered whether predictions were less accurate for pairs with pseudowords as a consequence of poorer actual memory (Thomas et al., 2011). Moreover, the effect of meaningfulness on metacognitive bias might result from a reduced familiarity with pseudowords than words as they were never encountered before the experiment. In Experiments 6.2 and 6.3, we aimed at exploring such hypotheses.

### 6.3. Experiment 6.2: Improving processing of meaningless items

We aimed at better understanding results obtained in Experiment 6.1 by allowing more time for encoding of meaningless items for half of the participants (see Thomas et al., 2011 for an example of time manipulation as a means for decreasing discrepancy in performance). By doing so, we wanted to check if PoR accuracy is reduced when semantic processing is not possible simply because meaningless items are harder to process at encoding. The alternative is that even when meaningless items are given more time at encoding, PoR accuracy is still impacted by the lack of access to semantic features. Therefore, the goal was to reduce discrepancies in first-order performance between the different conditions. To do so, we compared a group in which meaningless items were given more time to be processed than meaningful items and a group which saw both types of stimuli for the same amount of time as in Experiment 6.1.

We expected that increasing presentation time of pseudowords would improve recognition and thus reduce the discrepancy in memory performance between meaningless and meaningful targets as well as the difference between targets cued by meaningful and meaningless targets. Based on Experiment 6.1, we expected participants to make higher PoR for pairs with a word cue, in both groups. If the magnitude of PoR was based on first-order performance, PoR should be higher in conditions helped by the increase of presentation time. However, if predictions were based on the nature of the cue, we did not expect that increasing the encoding time of pseudowords would change the magnitude of the judgments. Moreover, if the effect of target-type on sensitivity found in the first study resulted from lack of time to process pseudowords which led participants to inaccurately predict their performance, giving more time to process pseudowords should decrease the target-type effect on sensitivity. Therefore, the target-type effect should have been lower in the group with increased encoding time for pseudowords. However, if this effect was due to lack of semantic processing of pseudowords beyond reading time constraints at encoding, we did not expect any difference in the effects on metacognitive sensitivity between the two groups.

#### 6.3.1. Method

##### ***Participants***

A hundred and thirty participants were recruited via the Prolific platform (<https://www.prolific.co/>). All were more than 18 years old and native French speakers and reported having no literacy difficulties on the platform ( $M_{age} = 26.64$ ;  $SD_{age} = 5.60$ ; 79 males). The sample size was determined based on the effect size of the nature of the target on gamma correlations observed in Experiment 6.1 ( $d_z = 0.57$ ), a power of .99 and considering around 15% of our sample was excluded in Experiment 6.1. Data from 45 participants were excluded

from the analyses (see [Results](#) section for more details). Therefore, data from 85 participants were included in the analyses (39 in the control group and 46 in the experimental group;  $M_{\text{age}} = 26.33$ ;  $SD_{\text{age}} = 5.48$ ; 55 males). All participants provided informed consent.

### **Material**

We used the same words and pseudowords as in Experiment 6.1. A few pseudowords were changed to increase the frequency of trigrams. Cues and targets were counterbalanced. The task was run on the Gorilla platform (<https://app.gorilla.sc/>).

### **Procedure**

A procedure similar to Experiment 6.1 was used: participants performed 2 blocks of the same task with different items. There was no distractor task in this experiment. Each block was split into 2 parts: learning of 60 cue-target pairs and PoR/recognition phase. In Experiment 6.1, cues and targets were presented simultaneously, but in this experiment, cues were always presented alone first and targets appeared after some time with the cues still present. Participants were assigned to one of two groups. In one group (i.e., control group), all the pairs were seen for 2s in total (as in Experiment 6.1): cues were presented for 1s alone and targets appeared for an additional 1s. In the experimental group (i.e., time-manipulation, TM group), the presentation time of the cues and the targets depended on the nature of the stimuli: pseudowords were presented for 2s whereas words were presented for 1s. As such, in the TM group PW/PW pairs were presented for 4s, W/PW and PW/W for 3s, and W/W pairs for 2s in total. Predictions were made on the same 5-point Likert scale as in Experiment 6.1.

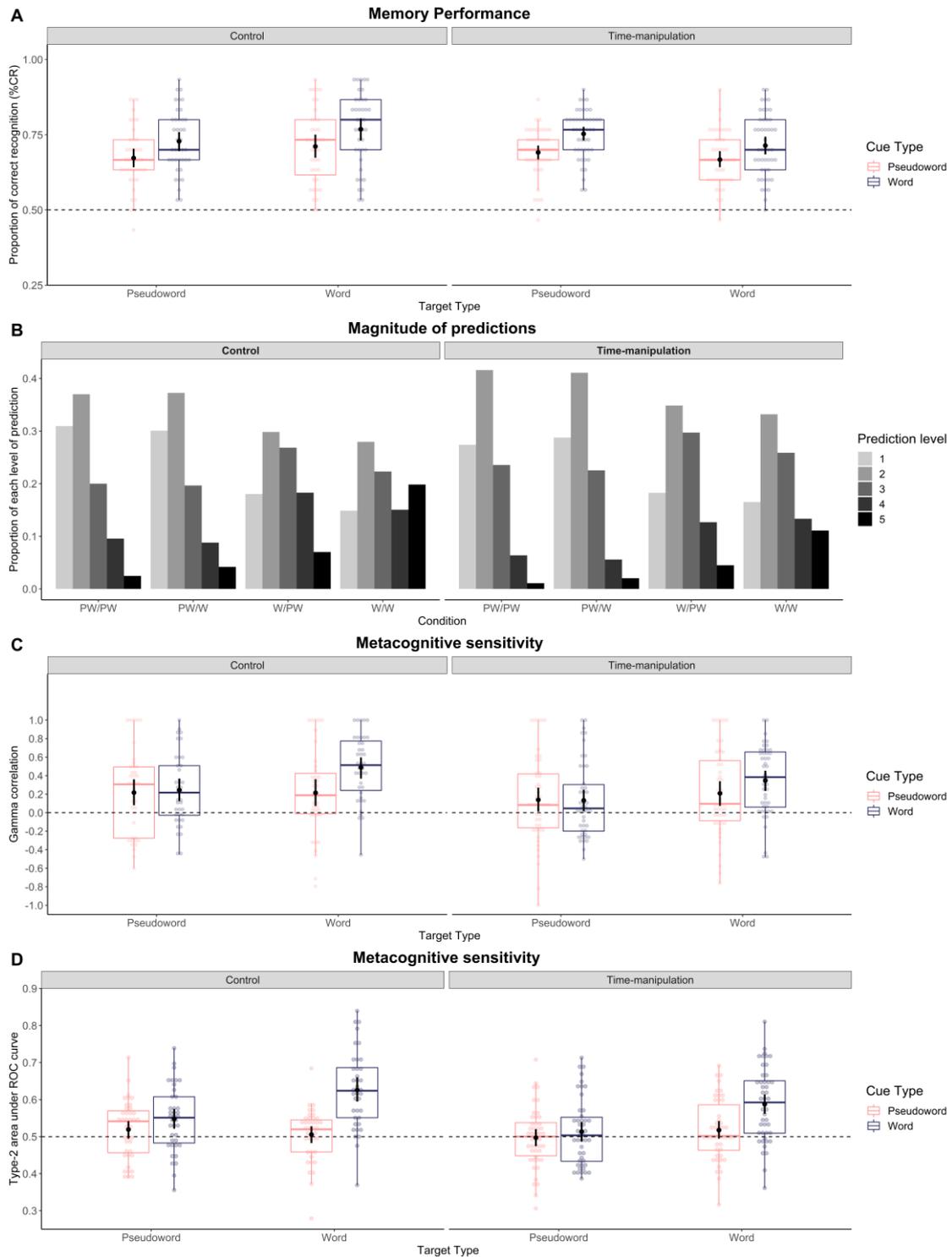
**Analyses.** The same analyses as in Experiment 6.1 were performed with the group as an additional variable. If not indicated otherwise, analyses met normality prerequisites for the use of linear models. To facilitate reading, we report statistics using linear models in text even when prerequisites were not met, but alternate statistics are reported as footnotes. Analyses were run on Rstudio with an error probability set at  $\alpha = .05$ .

## **6.3.2. Results**

### **Data exclusion**

As preregistered, data from 27 participants were excluded because they performed at chance in at least one block (13 in the control group and 14 in the TM group). We also excluded data from 11 participants who performed above 95%CR in at least one condition (7 in the control group and 4 in the TM group). We excluded 4 more participants from the analyses because of a lack of variation in their PoRs (4 from the control group). Finally, 3 other participants were

removed from the analyses as they presented extreme values ( $> 3$  SD from the mean) in at least one condition on one measure.



**Figure 6.3** : Results of Experiment 6.2.

A) First-order performance, B) Magnitude of judgments, C) Gamma scores, and D) AUROC2 scores.

Each dot represents one participant. Black dots and black lines are the means and the 95% confidence interval around it. For metacognitive bias, scores below zero suggest underestimation of performance and scores above zero suggest overestimation.

### **Memory performance**

Here we report statistics using %CR as the variable but the same pattern of results was found using  $d'$  as a measure of performance. Means and standard deviations for each condition and group are reported in [Appendix F](#). Memory performance did not significantly differ between groups ( $F(1,83) = 0.72, p = .40, \eta_p^2 = 0.009 [0, 0.09]$ )<sup>8</sup>, see Figure 6.3a. The effect of target-type on memory was not significant ( $F(1,83) = 0.15, p = .70, \eta_p^2 = 0.002 [0, 0.06]$ ) but interacted with the group ( $F(1,83) = 12.32, p < .001, \eta_p^2 = 0.13 [0.03, 0.27]$ ). In the control group, words were better recognized than pseudowords ( $t(38) = 2.32, p = .03, d_z = 0.37 [0.04, 0.70]$ ), whereas in the TM group pseudowords were better recognized than words ( $t(45) = 2.68, p = .01, d_z = 0.40 [0.09, 0.70]$ ). The effect of cue-type on memory performance was significant ( $F(1,83) = 42.29, p < .001, \eta_p^2 = 0.34 [0.18, 0.48]$ ) but did not significantly interact with group ( $F(1,83) = 0.02, p = .89, \eta_p^2 < 0.001 [0, 0.04]$ ). Participants recognized more targets cued by words than targets cued by pseudowords. The nature of the target did not interact with the nature of its cue ( $F(1,83) = 0.23, p = .63, \eta_p^2 = 0.003 [0, 0.07]$ ). Finally, the three-way interaction (target-type\*cue-type\*group) was not significant ( $F(1,83) = 0.36, p = .55, \eta_p^2 = 0.004 [0, 0.07]$ ).

Of note in this analysis, there was only one interaction with group, for the effect of target type. Our aim was to manipulate the difference in performance between the words and the pseudowords, and whilst we had no effect on the performance according to the type of cue used, we found differences according to group for the type of target. As in Experiment 6.1, in the control group, performance was higher for words than pseudowords, but this pattern was reversed in our group with extra study time.

### **Magnitude of judgments**

The magnitude of PoR was not significantly different between the two groups ( $F(1,83) = 1.27, p = .26, \eta_p^2 = 0.02 [0, 0.10]$ )<sup>9</sup>, see [Appendix F](#) for the means. The target-type effect was significant: PoR were lower for pseudoword targets than for word targets ( $F(1,83) = 39.76, p < .001, \eta_p^2 = 0.32 [0.17, 0.46]$ )<sup>10</sup> and this effect interacted with group ( $F(1,83) = 4.10, p = .046, \eta_p^2 = 0.05 [0, 0.16]$ )<sup>11</sup>: the target-type effect was slightly larger in the control group ( $t(38) = 4.89, p < .001, d_z = 0.78 [0.43, 1.15]$ )<sup>12</sup> than in the TM group ( $t(45) = 3.74, p < .001, d_z = 0.55 [0.24, 0.87]$ ). Participants also made lower PoR when cues were pseudowords than when cues were words;  $F(1,83) = 123.21, p < .001, \eta_p^2 = 0.60 [0.47, 0.69]$ . This effect did not interact with the group ( $F(1,83) = 2.45, p = .12, \eta_p^2 = 0.03 [0, 0.13]$ ). The nature of the target significantly interacted with the nature of the cue ( $F(1,83) = 30.40, p < .001, \eta_p^2 = 0.27 [0.12, 0.41]$ ). PoR

<sup>8</sup> Levene's test revealed heteroscedasticity ( $p = 0.04$ ). Welch t-test :  $t_{68.097} = 0.83, p = .41$ .

<sup>9</sup> Levene's test revealed heteroscedasticity ( $p = 0.046$ ). Welch t-test :  $t_{70.66} = 1.10, p = .27$ .

<sup>10</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 3025, p < .001$ .

<sup>11</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $W = 686, p = .06$ .

<sup>12</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $W = 683.5, p < .001$ .

were influenced by the nature of the target only when the cue was a word ( $t(84) = 6.99, p < .001, d_z = 0.76 [0.52, 1.00]$ )<sup>13</sup> but not when the cue was a pseudoword ( $t(84) = 0.53, p = .60, d_z = 0.06 [-0.16, 0.27]$ )<sup>14</sup>. The three-way interaction (cue-type\*target-type\*group) was not significant ( $F(1,83) = 0.56, p = .46, \eta_p^2 = 0.007 [0, 0.08]$ ). See Figure 6.3b for proportion of each judgment given.

In sum, overall the groups' performance did not differ, and nor did the magnitude of their predictions. Whereas the effect of target type was not significant for performance, it was for the magnitude of predictions; and this effect was larger in the control group. The effect of cue type was significant for performance and for the magnitude of people's predictions. In only one case was there an interaction with group – the difference between predictions according to target type was larger in the control group.

### ***Metacognitive sensitivity and efficiency***

*Gamma correlations.* Gamma correlations were not different between groups ( $F(1,83) = 2.33, p = .13, \eta_p^2 = 0.03 [0, 0.13]$ , see Figure 6.3c). Analyses revealed significant main effects of target-type ( $F(1,83) = 7.94, p = .006, \eta_p^2 = 0.09 [0.01, 0.22]$ ) and cue-type ( $F(1,83) = 11.06, p = .001, \eta_p^2 = 0.12 [0.02, 0.26]$ ) on gamma correlations. Gamma scores were higher for pairs with a word target than for pairs with a pseudoword target, and were also higher for pairs with a word cue than for pairs with a pseudoword cue. Neither of these effects interacted with the group (target-type\*group:  $F(1,83) = 0.05, p = 0.82, \eta_p^2 < .001 [0, 0.05]$ ; cue-type\*group:  $F(1,83) = 1.67, p = .20, \eta_p^2 = 0.02 [0, 0.11]$ ). However, the interaction between target-type and cue-type was also significant ( $F(1,83) = 4.49, p = .037, \eta_p^2 = 0.05 [0, 0.17]$ ): pairs with a word cue produced higher gamma correlations when the target was also a word than when it was a pseudoword;  $t(84) = 4.62, p < .001, d_z = 0.59 [0.36, 0.83]$ <sup>15</sup>, whereas no such difference was found for pairs with a pseudoword cue ( $t(84) = 0.47, p = .64, d_z = 0.05 [-0.16, 0.26]$ ). The three-way interaction (target-type\*cue-type\*group) was not significant ( $F(1,83) = 0.31, p = .58, \eta_p^2 = .004 [0, 0.07]$ ). In short there was no group effect and no interactions with group. The effect of target type and the interaction between cue type and target type match those in Experiment 6.1, but this time, we also have a significant effect of cue type on the gamma correlation.

*Type-2 ROC curves.* The same pattern of results as gamma scores analyses was found when using AUROC2 as a measure of metacognitive sensitivity except for the addition of a cue-type by group interaction. Analyses conducted on AUROC2 revealed significant main effects of

<sup>13</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 2788.5, p < .001$ .

<sup>14</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 1711.5, p = .66$ .

<sup>15</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 2772, p < .001$ .

target-type ( $F(1,83) = 18.34, p < .001, \eta_p^2 = 0.18 [0.06, 0.33]$ ) and cue-type ( $F(1,83) = 64.14, p < .001, \eta_p^2 = 0.44 [0.28, 0.56]$ ), as well as an interaction of the target-type with cue-type ( $F(1,83) = 15.72, p < .001, \eta_p^2 = 0.16 [0.04, 0.30]$ ). AUROC2 were higher for pairs with word targets than for pairs with pseudoword targets (see Figure 6.3d). However, simple effects analyses revealed that this difference was only significant when the cue was also a word ( $t(84) = 6.07, p < .001, d_z = 0.66 [0.42, 0.90]$ ) but not when the cue was a pseudoword ( $t(84) = 0.35, p = .73, d_z = 0.04 [-0.18, 0.25]$ ). Participants also had higher AUROC2 for pairs with a word cue than for pairs with a pseudoword cue but the effect of cue-type slightly interacted with the group ( $F(1,83) = 4.54, p = 0.036, \eta_p^2 = 0.05 [0, 0.17]$ ) which was not the case using gamma correlations as a variable. This effect was larger in the control group ( $t(39) = 7.03, p < .001, d_z = 1.13 [0.73, 1.54]$ ) than in the TM group ( $t(45) = 4.27, p < .001, d_z = 0.63 [0.31, 0.95]$ ). Similar to analyses on gamma correlations, we did not find any significant effect of group ( $F(1,83) = 3.15, p = .08, \eta_p^2 = 0.04 [0, 0.14]$ ), interaction of target-type\*group ( $F(1,83) = 0.71, p = .40, \eta_p^2 = 0.008 [0, 0.09]$ ), nor double interaction target-type\*cue-type\*group ( $F(1,83) = 1.05, p = .31, \eta_p^2 = 0.01 [0, 0.10]$ ) in AUROC2 analyses. The results of this analysis are very similar to that presented in Experiment 6.1, with the addition of one factor which changes when comparing groups: the response to a cue type. This showed that there was a larger benefit of words cues to prediction sensitivity in the control condition than the experimental condition.

*Metacognitive efficiency.* As preregistered, we used the Hmeta-d bayesian framework to estimate M-ratio as it is the best fitted measure for designs with less than 100 trials. However, as for Experiment 6.1, the MCMC chains for estimating the effects of cue-type and target-type on M-ratio did not converge. Therefore, we do not report these results here. Graphs of the models' outputs can be found in [Appendix I](#).

### 6.3.3. Interim Discussion

Increasing presentation time of meaningless items impacted actual recognition performance as expected but performance across all four conditions was still not equated. Participants who were allowed more time to learn meaningless targets later recognized them better than meaningful ones. On the other hand, participants who saw meaningless and meaningful items for the same amount of time at learning recognized more meaningful than meaningless targets, consistent with results from Experiment 6.1. Note that usually, better first-order performance is associated with better metacognitive sensitivity (Fleming & Lau, 2014; Vuorre & Metcalfe, 2021), but according to our hypothesis about meaningfulness aiding metacognitive predictions, we expected worse metacognitive sensitivity for pairs with meaningless items. Therefore, better recognition of pseudowords than words in the time-

manipulation group is less of a problem for interpretation of the effects on metacognitive measures than an effect in the opposite direction (as it was the case in Experiment 6.1). However, time-manipulation did not significantly change the cue-type effect on recognition, which was equivalent in both groups, as targets associated with meaningless cues were less recognized than targets with meaningful cues.

Consistent with Experiment 6.1 and the hypothesis of a combined implication of cue-familiarity and accessibility to target (Koriat & Levy-Sadot, 2001), meaningfulness of cues affected the magnitude of the PoR, and when the cue was meaningful, PoR also changed as a function of the type of target. Participants in both groups predicted worse recognition when they saw a meaningless cue. When the cue was meaningful, they also predicted less recognition of meaningless items than meaningful items, even participants who were given more time to process such items despite them recognizing more meaningless targets. This result was the same as in the control group (even though it was very slightly reduced) and in Experiment 6.1. Taken together, these results show that even when more time is given for processing of meaningless items and when actual recognition memory is increased, participants are still affected by the type of cue and the type of target when the cue is familiar when making PoR.

Finally, we replicated the effect of meaningfulness of the target on metacognitive sensitivity. Even when meaningless items were presented for a longer time at learning, participants were still less able to accurately predict their performance when at least one item of the pair was meaningless. Note that time manipulation slightly reduced the effect of the meaningfulness of the cue but only when using AUROC2 as a measure and the effect was still significant ( $p < .001$ ). These results show that even when meaningless items are given more opportunity for being processed (without semantic information) and when meaningless targets are better remembered, the lack of semantic information still seems to prevent participants from accurately predicting their performance. We will discuss these results more thoroughly in the General Discussion.

#### **6.4. Experiment 6.3: Increasing familiarity with meaningless items**

The goal of this third experiment was to facilitate encoding of pseudowords by increasing their familiarity by prior exposure before the pair-learning phase. Predictions are greatly influenced by cue familiarity (e.g., Metcalfe et al., 1993; Reder & Ritter, 1992) and pseudowords are by nature less familiar than words because they have never been encountered before. Completely equalizing “initial” familiarity (i.e., how familiar the person is with a word before learning the pairs based on experience) between meaningful and meaningless items is hard, if not impossible due to the intrinsic nature of what a pseudoword is. However, one way of increasing familiarity is to manipulate the number of times an item is presented (e.g., Metcalfe et al., 1993).

Here, we wanted to improve recognition of meaningless items to reduce the possible influence of first-order performance on predictions. As PoR are greatly influenced by familiarity with the cue, we also aimed at increasing “initial” familiarity of meaningless items. As such, we created two groups. One group phonologically processed the meaningless items of the pairs before the encoding phase (i.e., familiarization phase). The other group performed a filler task. By doing so, we wanted to investigate how this familiarization with meaningless cues and targets would affect PoR accuracy, and see whether the meaninglessness of cues and targets still influenced metacognition sensitivity under such conditions.

Based on Experiment 6.1 and Experiment 6.2 results, we expected better recognition of word targets than pseudoword targets in the control group. We expected this effect to be lower or reversed in the group who were familiarized with pseudowords. Because participants read pseudowords before the learning in one group and suspected that it would help processing of such items at encoding, we also expected the cue-type effect on recognition performance to be lower in this group compared to the control group. As prediction magnitude is usually influenced by cue familiarity (e.g., Schwartz & Metcalfe, 1992) and pseudowords are by nature less familiar than words, we expected lower PoR and greater underestimation for pairs with a pseudoword cue than for pairs with a word cue. However, this effect should therefore be reduced in the familiarized group. Finally, even if read before learning, less semantic information about the target is accessible when it is meaningless than when it is meaningful and metacognitive sensitivity is influenced by partial information that can be retrieved about the target (Thomas et al., 2012), even though participants can still rely on other types of information (e.g., phonological/orthographic cues, episodic details; Koriat, 1993; Schwartz et al., 2014). Therefore, we expected a main effect of the target on metacognitive sensitivity. Moreover, as for Experiment 6.1 and 6.2, we expected metacognitive sensitivity to be the best when both the cue and the target are meaningful.

### 6.4.1. Method

#### **Participants**

We recruited 165 French-speaking participants<sup>16</sup> on the Prolific platform. All were above 18 years old with no language-related disorder ( $M_{age} = 25.73$ ;  $SD_{age} = 5.11$ ; 87 males). As for Experiment 6.2, the initial sample size ( $n = 130$ ) was determined considering the effect size of target type on gamma scores in Experiment 6.1 ( $d_z = 0.58$ ), a power of .99 and taking into account that 15% of participants were removed from analyses in Experiment 6.1. Forty-seven participants (28 in the control group, 19 in the experimental group) were excluded from the analyses because they did not meet inclusion criteria (see [Results](#) section for details). As such, 118 participants ( $M_{age} = 25.95$ ;  $SD_{age} = 5.21$ ; 59 males) were included in the analyses (60 in control group and 58 in experimental group). Participants gave informed consent before study completion.

#### **Material**

We used the same material as in Experiment 6.2 (180 words and 180 pseudowords).

#### **Procedure**

Participants completed 2 blocks of 3 phases. For the first phase, we created two groups of participants. One group (familiarization, Fam-group) completed a priming phase in which they saw all the future pseudowords targets and cues one by one (in a random order) and for each, they were given 4s to say whether upper and lower lips had to touch to pronounce them. They did not see any words nor future distractors in this phase. This task ensured participants read all pseudowords before the experiment. In a control group, participants completed a filler task in which they had to count the number of corners in geometrical figures and say if they were more or less than 4. The other two phases of this experiment were similar to the learning (60 cue-target pairs) and PoR/recognition phases of Experiment 6.1. As in Experiment 6.1, each cue was simultaneously presented with its target for 2s.

#### **Analyses**

All analyses were the same as for Experiment 6.2. Analyses were run on Rstudio with an error probability set at  $\alpha = .05$ .

---

<sup>16</sup> We initially recruited 130 participants as pre-registered. However, counterbalancing was not successful due to a programming error. Therefore, we recruited 35 additional participants to have the correct counterbalancing of cues and targets.

## 6.4.2. Results

### **Data exclusion**

We excluded data from 23 participants because of floor performance (<55%CR) in at least one block (16 in control group, 7 in Fam-group), 9 because of ceiling performance in at least one condition (3 in control group and 6 in Fam-group), and 10 more participants with not enough variation in levels of PoR. Data from 5 additional participants (3 from control group and 2 from Fam-group) were removed from analyses as they presented extreme values (> 3 SD from the mean) on at least one measure.

### **Memory performance**

We report statistics of %CR here but we found a similar pattern of results using  $d'$  as a measure of performance (see [Appendix G](#) for the details of means). The Fam-group performed slightly better than the control group on the recognition test ( $F(1,116) = 4.17, p = .043, \eta_p^2 = 0.03$  [0, 0.12], see Figure 6.4a). Overall, memory performance was not affected by target-type ( $F(1,116) = 1.73, p = .19, \eta_p^2 = 0.01$  [0, 0.08]) but this effect of target-type differed between groups ( $F(1,116) = 8.18, p = 0.005, \eta_p^2 = 0.07$  [0.01, 0.17]). Indeed, participants in the Fam-group recognized more pseudoword targets than word targets ( $t(57) = 3.14, p = .003, d_z = 0.41$  [0.14, 0.68]) but such difference was not found in the control group ( $t(59) = 1.04, p = .30, d_z = 0.13$  [-0.12, 0.39]). Analyses also revealed a significant effect of cue-type ( $F(1,116) = 40.35, p < .001, \eta_p^2 = 0.26$  [0.13, 0.38]) as targets cued by words were better recognized than targets cued by pseudowords. This effect did not interact with the group ( $F(1,116) = 0.008, p = .93, \eta_p^2 < 0.001$  [0, 0.02]). Finally, no significant interaction target-type\*cue-type was found ( $F(1, 116) = 0.04, p = .84, \eta_p^2 < 0.001$  [0, 0.03]) nor target-type\*cue-type\*group ( $F(1, 116) = 0.82, p = .37, \eta_p^2 = 0.007$  [0, 0.07]).

### **Magnitude of judgments**

The magnitude of PoR did not differ between groups ( $F(1, 116) = 0.86, p = 0.36, \eta_p^2 = 0.007$  [0, 0.07]). Participants predicted better memory performance for word targets than pseudoword targets ( $F(1, 116) = 48.01, p < .001, \eta_p^2 = 0.29$  [0.16, 0.41])<sup>17</sup> and no significant interaction with group was found ( $F(1, 116) = 0.51, p = .48, \eta_p^2 = 0.004$  [0, 0.06])<sup>18</sup>. Participants also made higher PoR when the cue was a word than when it was a pseudoword ( $F(1, 116) = 90.87, p < .001, \eta_p^2 = 0.44$  [0.31, 0.55])<sup>19</sup> and this effect did not interact with the group ( $F(1, 116) = 0.19, p = .67, \eta_p^2 = 0.002$  [0, 0.05])<sup>20</sup>. However, the effect of target-type on magnitude of judgments

<sup>17</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 5439, p < .001$ .

<sup>18</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 1705.5, p = .85$ .

<sup>19</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 6373, p < .001$ .

<sup>20</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 1672.5, p = .72$ .

depended on the type of cue ( $F(1, 116) = 46.75, p < .001, \eta_p^2 = 0.29 [0.16, 0.41]$ )<sup>21</sup> as this effect was only significant when the cue was a word ( $t(117) = 7.80, p < .001, d_z = 0.72 [0.52, 0.92]$ )<sup>22</sup> but not when it was a pseudoword ( $t(117) = 1.27, p = .21, d_z = 0.12 [-0.06, 0.30]$ ). When the cue was a word, participants made higher PoR for the recognition of word targets than that of pseudoword targets. The cue-type\*target-type interaction was not different between groups ( $F(1, 116) = 0.004, p = .95, \eta_p^2 < 0.001 [0, 0.01]$ )<sup>23</sup>. In short, in this experiment, familiarity increased recognition performance for the pseudowords, and in the Fam-group, the pseudoword targets were better recognized than the words. However, predictions of performance are such that people predicted higher performance for words than pseudowords, something that is the same in both groups. See Figure 6.4b for proportion of each judgements.

### **Metacognitive sensitivity and efficiency**

*Gamma correlations.* Analyses conducted on gamma scores revealed a significant effect of the target-type ( $F(1, 116) = 38.28, p < .001, \eta_p^2 = 0.25 [0.12, 0.37]$ ) as well as an interaction between target-type and cue-type ( $F(1, 116) = 11.65, p < .001, \eta_p^2 = 0.09 [0.02, 0.20]$ , see Figure 6.4c). Gamma correlations were higher for pairs with word targets than for pairs with pseudoword targets and this difference was only significant when the cue was also a word ( $t(117) = 6.73, p < .001, d_z = 0.62 [0.42, 0.82]$ ) but not when it was a pseudoword ( $t(117) = 1.67, p = .09, d_z = 0.16 [-0.03, 0.34]$ ). None of these effects interacted with group (target-type\*group :  $F(1, 116) = 0.25, p = .62, \eta_p^2 = 0.002 [0, 0.05]$ ; target-type\*cue-type\*group:  $F(1, 116) = 0.24, p = .63, \eta_p^2 = .002 [0, 0.05]$ ). None of the other effects reached significance (group :  $F(1, 116) = 0.13, p = .72, \eta_p^2 = 0.001 [0, 0.04]$ )<sup>24</sup>; cue-type:  $F(1, 116) = 1.98, p = .16, \eta_p^2 = 0.02 [0, 0.09]$ ; cue-type\*group:  $F(1, 116) = 0.02, p = .89, \eta_p^2 < 0.001 [0, 0.03]$ ).

*Type 2 ROC curves.* The analyses conducted on AUROC2 revealed a similar pattern of results except that the type of cue had a significant effect ( $F(1, 116) = 26.61, p < .001, \eta_p^2 = 0.19 [0.08, 0.31]$ ) which was not the case with gamma correlations (see Figure 6.4d). AUROC2 were larger for pairs with a word cue than for pairs with a pseudoword cue, and this effect did not interact with the group ( $F(1, 116) = 0.03, p = 0.87, \eta_p^2 < 0.001 [0, 0.03]$ ). Like gamma scores, AUROC were higher when the target was a word than when it was a pseudoword ( $F(1, 116) = 70.04, p < .001, \eta_p^2 = 0.38 [0.24, 0.49]$ ) and this effect did not interact with the group ( $F(1, 116) = 0.16, p = .69, \eta_p^2 = 0.001 [0, 0.04]$ ). However, as for gamma correlations, this effect of target-type was higher when the cue was a word ( $t(117) = 9.41, p < .001, d_z = 0.87 [0.66, 1.08]$ ) than when it was a pseudoword ( $t(117) = 2.50, p = 0.014, d_z = 0.23 [0.05, 0.41]$ );  $F(1, 116) = 24.40, p <$

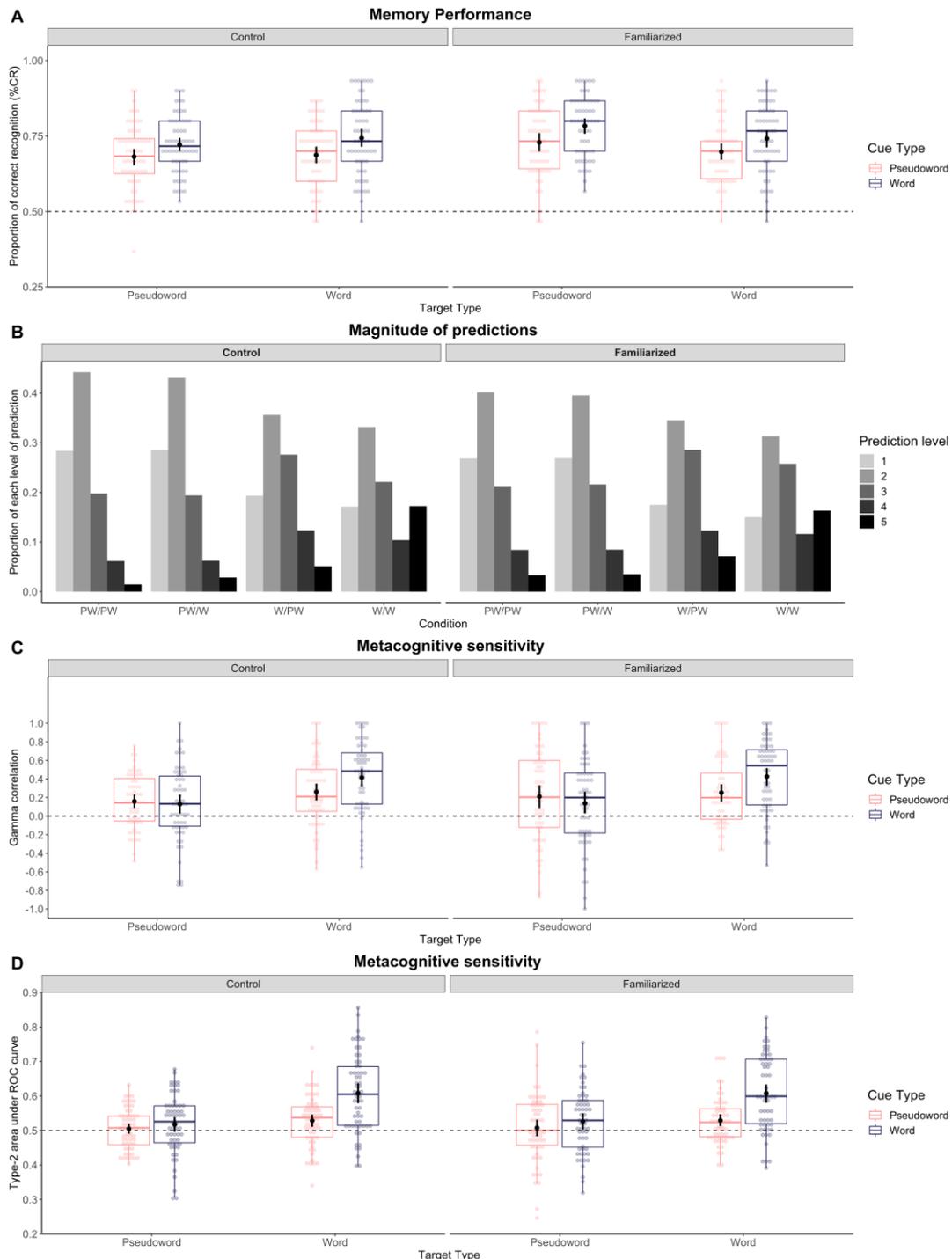
<sup>21</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 5333, p < .001$ .

<sup>22</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 5981, p < .001$ .

<sup>23</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 1822, p = .66$ .

<sup>24</sup> The residuals of the model were not normally distributed. Mann-Whitney :  $W = 1801, p = .74$ .

.001,  $\eta_p^2 = 0.17$  [0.07, 0.30]. We did not find any significant effect of group on AUROC2 ( $F(1, 116) = 0.09, p = 0.77, \eta_p^2 < 0.001$  [0, 0.04])<sup>25</sup>, and no target-type\*cue-type\*group interaction ( $F(1, 116) = 0.06, p = 0.81, \eta_p^2 < 0.001$  [0, 0.03]).



**Figure 6.4 :** Results of Experiment 6.3.

A) First-order performance, B) Magnitude of judgments, C) Gamma scores, and D) AUROC2 scores. Each dot represents one participant. Black dots and black lines are the means and the 95% confidence interval around it. For metacognitive bias, scores below zero suggest underestimation of performance and scores above zero suggest overestimation.

<sup>25</sup> The residuals of the model were not normally distributed. Mann-Whitney :  $W = 1723, p = .93$ .

*Metacognitive efficiency.* We tried to estimate the effects of cue-type, target-type and their interaction on M-ratio using hierarchical models but as for the other experiments, the models did not converge. Graphs of the output are provided in [Appendix I](#).

### 6.4.3. Interim discussion

In this experiment, we improved recognition in one group by increasing familiarity of the meaningless targets. In both groups, participants still recognized more targets cued by meaningful cues than by meaningless cues. Therefore, as for Experiment 2, performance was not perfectly equated between conditions, but our manipulation improved memory of meaningless targets. Note that the effect size of prior presentation of meaningless stimuli on memory performance might be underestimated as less participants were removed from the analyses because of floor performance in the Fam-group ( $n = 7$ ) than in the control group ( $n = 16$ ).

Exposing participants to meaningless items before task completion did not change results on predictions' magnitude. Participants were still affected by the type of cue when making predictions as they predicted worse recognition when seeing a meaningless cue than a meaningful cue, in line with the cue-familiarity hypothesis (Schwartz & Metcalfe, 1992). As we did not measure familiarity with the cues per se, we cannot attest that prior exposure actually increased familiarity with meaningless items in the Fam-group. Therefore, multiple hypotheses can be advanced. For one, familiarity with meaningless cues may not have been increased sufficiently by prior exposure: prior exposure increased familiarity of meaningless items but they were still less familiar than real words. Consistent with results from Experiments 6.1 and 6.2, when the cue was meaningful, participants predicted poorer recognition of meaningless targets than meaningful targets.

Participants gave similar PoR in both groups despite memory performance for meaningless items being increased in the Fam-group (i.e., increased bias). This might be a consequence of using a between-subject manipulation (see Carroll & Nelson, 1993 for an example of effects that are more detectable in within-subject designs than in between-subject designs). Participants might use the same parts of the scale and change their judgments the same way based on within-subject manipulations. As in Experiment 6.1 and 6.2, participants of both groups were less metacognitively sensitive for pairs in which at least one item was meaningless. This shows that even when participants are better at recognizing meaningless targets, participants are still better at predicting memory for meaningful targets if the cue is meaningful. This result will be discussed further in the [General Discussion](#).

### 6.5. Experiment 6.4: The role of recall

The goal of this last experiment was to assess the role of retrieval and recall abilities in the influence of semantic processing on PoR accuracy. The PoR paradigm used in the experiments described should rely on associative learning and we find interesting results according to the characteristics of the cue in the above experiments. However, paired associates can be learnt with different strategies. For example, making an interactive image of both items of the pair or creating a meaningful sentence including both items improve recall and recognition of pairs better than only repeating or silently rehearsing them (Bower & Winzenz, 1970; Straub & Granaas, 1992). These strategies can be spontaneously used for pairs of meaningful items, especially concrete words (e.g., mental imagery, sentence generation). However, only limited and less efficient strategies (e.g., repetition, phonological or orthographic association) can be used for memorizing pairs with at least one meaningless item. This intrinsic difference in the opportunity of using such strategies alone cannot explain the results on the recognition test in Experiments 6.1, 6.2, and 6.3. Indeed, even though words were better recognized than pseudowords, especially when they were cued by a word when all types of items were presented for the same amount of time, recognition performance for pseudowords was increased with longer presentation time at encoding. However, even if participants were not explicitly asked to *recall* the target in Experiments 6.1, 6.2 and 6.3, it is possible that when making their PoR they were able to spontaneously retrieve it in some trials, especially those containing both meaningful cues and targets. Previous work suggests that feeling-of-knowing (FOK) sensitivity is better when FOK are made on all trials rather than only on trials with failed recall (Schwartz et al., 2016).

However, in the foregoing experiments it is possible that participants do not tackle the task as they would in a pair-learning task, since the final test is a 2AFC recognition task. Whilst the sensitivity and magnitude of people's judgements makes it clear that the cue is processed as being relevant for the task (i.e., it is not ignored completely), the final recognition task can be solved by familiarity alone: which of two words is more familiar. This explains the high level of performance in Experiment 6.3: one of the pseudowords was more familiar than the other. That is, we assumed that at point of making the metacognitive prediction that participants attempted recall and this informed their judgements. However, despite the characteristics of the cue influencing the predictions made and the performance, it is not known whether participants did indeed use this strategy. It is possible also that participants did attempt recall but that it was misleading, as is borne out in the results of Experiment 6.3: we might interpret that participants in the familiarity group were not aware of how easy the 2AFC with one more familiar target might be.

In this experiment 6.4, we aimed at exploring whether the better sensitivity for meaningful pairs compared to pairs with at least one meaningless item could be explained by increased target retrievability when making the PoR. According to the memory constraint hypothesis (Hertzog et al., 2010), metacognitive sensitivity is limited when there is low memory strength. Our hypothesis is that participants are better able to recall the target when it is meaningful and when it is cued by a meaningful item because of more efficient associative learning (and thus, greater memory strength) and that this leads participants to be more metacognitively sensitive. We used a similar paradigm as in the time-manipulation group from Experiment 6.2 as it was the procedure which led to the lowest difference in first-order performance between conditions, but we added a recall phase before the PoR, similar to a standard FOK paradigm. As successful recall should be highly predictive of correct recognition, including trials with correct recall in the analyses should lead to increased metacognitive sensitivity compared to when only including trials with unsuccessful recall are included. Therefore, we ran analyses on all trials and on trials with incorrect recall only and compared outcomes to evaluate whether considering recall performance might account for differences in metacognitive sensitivity between conditions.

We expected participants to recall more meaningful targets paired with meaningful cues than meaningless targets or than meaningful targets paired with a meaningless cue because of stronger associative learning. Based on results from Experiment 6.2, we expected equal recognition performance for meaningful and meaningless targets, or a small advantage for meaningless items (because they were presented for a bit more time). We also expected better recognition of targets learnt with meaningful cues. We expected higher PoR for pairs with meaningful cues than for pairs with meaningless cues. Importantly, as no semantic information about the target can be accessed and retrieval of partial information about the target might be harder when the item is meaningless, we expect greater sensitivity for meaningful targets compared to meaningless targets. We hypothesized that the effects observed on Experiments 6.1, 6.2, and 6.3 regarding sensitivity might result from increased access to targets in some conditions (especially the W-W condition) because of stronger associative learning. Therefore, we expect lower effect of target and lower target-type\*cue-type interaction when only trials with no recall are included in the analyses.

### **6.5.1. Method**

#### ***Participants***

We recruited 42 French speaking psychology students who were tested in the lab at the University Grenoble-Alpes ( $M_{age} = 19.64$ ,  $SD_{age} = 1.10$ , 6 men). All were above 18 years old and had normal or corrected to normal vision. None reported having neurodevelopmental,

neurological or psychiatric disorder. This sample size was determined on the effect size of target-type on gamma correlations in Experiment 6.1 ( $d_z = 0.57$ ), a power of .95 and  $\alpha = .05$ . However, 10 participants were removed from the analyses (see [Results](#) section for details). Therefore, 32 participants were included in the analyses ( $M_{age} = 19.56$ ,  $SD_{age} = 0.98$ , 6 men). All participants provided informed consent.

### **Material**

Pairs of words and pseudowords were the same as in Experiments 6.2 and 6.3. The experiment was run on E-prime 3 software (Psychology Software Tools, 2016).

### **Procedure**

Because increasing presentation-time of pseudowords at encoding increased memory performance for pseudoword targets in Experiment 6.2, we used a similar procedure here (to minimize differences in first-order performance between pseudowords and words). There was one change: participants were asked to recall the target before making a PoR. Therefore, all participants performed 2 blocks which consisted of 2 phases: 1) learning of 60 cue-target pairs, and 2) recall attempt, PoR and recognition. In the learning phase, cues were first presented alone before their target appeared with the cue still displayed. Pseudowords were displayed for 2s and words for 1s. During the recall phase, participants were given 15s to type the target. After the recall attempt, they made a PoR and were then presented with a 2AFC recognition test for each pair.

### **Analyses**

For the recall phase, we accepted as correct trials with typos, trials in which participants typed a pseudoword which was phonologically identical to the target (e.g., *bodin* instead of *baudain* which would be read the same in French), and trials in which pseudowords were one letter away from the correct target (e.g., *solat* instead of *socat*). We performed the same analyses as in Experiment 1, with an additional analysis on recall performance. To explore whether results from previous experiments resulted from difference in recall abilities, analyses of recognition performance and metacognitive measures were performed twice: once including all trials and once including only trials with no correct recall. All analyses were run on Rstudio with an error probability set at  $\alpha = .05$ . If not stated otherwise, data met normality prerequisites for the use of ANOVAs.

## 6.5.2. Results

### **Data exclusion**

We excluded 6 participants from the analyses: one performed below 55%CR in one block, four had ceiling performance in one condition (1 in PW/W condition and 3 in W/W condition), and one did not vary their PoR which prevented the calculation of gamma correlations. We excluded data from 4 more participants as they presented an extreme value ( $> 3$  SD from the mean) on at least one measure (3 on recall performance, 1 on magnitude).

### **Memory performance**

*Recall (proportion of correct recall).* Participants recalled more words than pseudowords ( $F(1, 31) = 37.77, p < .001, \eta_p^2 = 0.55 [0.30, .70]$ )<sup>26</sup> and more targets which were cued by a word ( $F(1, 31) = 35.06, p < .001, \eta_p^2 = 0.53 [0.28, 0.69]$ )<sup>27</sup> (see [Appendix H](#) for means). However, the effect of target type was influenced by cue type ( $F(1, 31) = 34.33, p < .001, \eta_p^2 = 0.53 [0.27, 0.69]$ )<sup>28</sup> as participants recalled significantly more words than pseudowords when the cue was also a word ( $t(31) = 6.07, p < .001, d_z = 1.07 [0.64, 1.53]$ )<sup>29</sup> with this effect being smaller when the cue was a pseudoword ( $t(31) = 2.33, p = 0.03, d_z = 0.41 [0.05, 0.78]$ )<sup>30</sup>. Participants were sometimes able to recall the target when both items of the pair were words ( $M_{W/W} = 0.17, SD_{W/W} = 0.16$ ) but very rarely or never in the other conditions ( $M_{PW/PW} = 0, SD_{PW/PW} = 0; M_{PW/W} = 0.009, SD_{PW/W} = 0.02; M_{W/PW} = 0.004, SD_{W/PW} = 0.01$ ).

When only considering trials with no correct recall, analyses were conducted on average on 30 ( $SD = 0$ ) trials in the PW/PW condition, 29.7 ( $SD = 0.68$ ) trials in the PW/W condition, 29.9 ( $SD = 0.34$ ) trials in the W/PW condition, and 24.9 ( $SD = 4.71$ ) trials in the W/W condition.

*Recognition (%CR and  $d'$ ).* When taking into account all trials, participants recognized more word targets than pseudoword targets ( $F(1, 31) = 4.95, p = .03, \eta_p^2 = 0.14 [0, 0.37]$ ) and more targets cued by words than targets cued by pseudowords ( $F(1, 31) = 13.10, p = .001, \eta_p^2 = 0.30 [0.06, 0.52]$ , see Figure 6.5a). The effect of target-type on recognition memory did not interact with cue-type ( $F(1, 31) = 0.56, p = .46, \eta_p^2 = 0.02 [0, 0.19]$ ). When considering only the trials with no correct recall, only the effect of cue-type was significant ( $F(1, 31) = 5.23, p = .03, \eta_p^2 = 0.14 [0, 0.37]$ ). Participants were better at recognizing targets cued by words than targets cued by pseudowords. No effect of target-type ( $F(1, 31) = 1.92, p = .18, \eta_p^2 = 0.06 [0, 0.27]$ ) nor a target-type\*cue-type interaction ( $F(1, 31) = 0.44, p = .51, \eta_p^2 = 0.01 [0, 0.18]$ ) were found. The same results were observed using  $d'$  as a measure of performance.

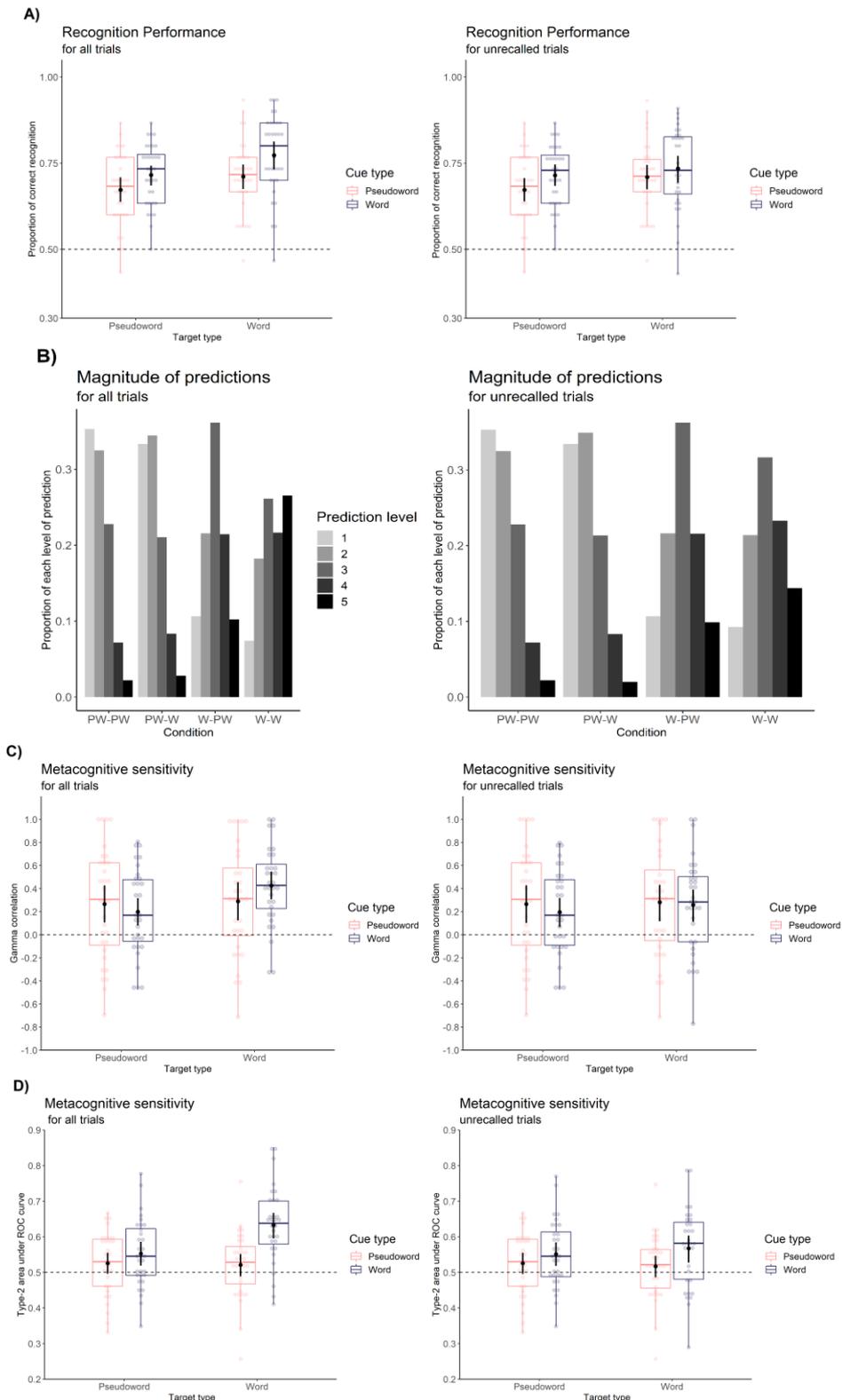
<sup>26</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 435, p < .001$ .

<sup>27</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 406, p < .001$ .

<sup>28</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 406, p < .001$ .

<sup>29</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 435, p < .001$ .

<sup>30</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 21, p = .03$ .



**Figure 6.5 : Results of Experiment 6.4.**

A) First-order performance, B) Magnitude of judgments, C) Gamma scores, and D) AUROC2 scores. Each dot represents one participant. Black dots and black lines are the means and the 95% confidence interval around it. For metacognitive bias, scores below zero suggest underestimation of performance and scores above zero suggest overestimation. Analyses were conducted on all trials (left) as well as on unrecalled trials only (right).

### ***Magnitude of judgments***

When taking into account all trials, participants made lower predictions of recognition of pseudowords than words ( $F(1, 31) = 29.96, p < .001, \eta_p^2 = 0.49 [0.23, 0.66]$ ) and lower predictions when the cue was a pseudoword than when it was a word ( $F(1, 31) = 110.67, p < .001, \eta_p^2 = 0.78 [0.63, 0.86]$ ), see Figure 6.5 for proportions. These findings are in keeping with how recognition varied according to these factors. However, the interaction was significant ( $F(1, 31) = 18.90, p < .001, \eta_p^2 = 0.38 [0.12, 0.58]$ ) with the effect of target-type being significant when the cue was a word ( $t(31) = 5.30, p < .001, d_z = 0.94 [0.52, 1.37]$ ) but not when it was a pseudoword ( $t(31) = 1.32, p = 0.20, d_z = 0.23 [-0.12, 0.59]$ ). When only trials with no correct recall were considered, the type of target ( $F(1, 31) = 6.12, p = .02, \eta_p^2 = 0.16 [0, 0.40]$ ) and the type of cue ( $F(1, 31) = 107.84, p < .001, \eta_p^2 = 0.78 [0.62, 0.86]$ ) still gave significant effects on the magnitude of judgments. Even when they were unable to recall the correct target, participants made lower judgments when they saw a pseudoword cue than when they saw a word cue. They also made lower predictions for pseudoword targets than for words. However, the target-type\*cue-type interaction was not significant when removing trials with correct recall from analyses ( $F(1, 31) = 3.26, p = .08, \eta_p^2 = 0.10 [0, 0.32]$ ).

### ***Metacognitive sensitivity and efficiency***

*Gamma correlations.*<sup>31</sup> First we considered responses for all trials. Gamma correlations were not significantly influenced by target-type ( $F(1, 31) = 3.41, p = .07, \eta_p^2 = 0.10 [0, 0.32]$ ) nor cue-type ( $F(1, 31) = 0.23, p = .63, \eta_p^2 = 0.007 [0, 0.16]$ , see Figure 6.5c). The interaction between effects of cue-type and target-type was not significant either ( $F(1, 31) = 1.69, p = 0.20, \eta_p^2 = 0.05 [0, 0.26]$ ). Moreover, none of these effects were significant when only looking at data from unrecalled trials (target-type:  $F(1, 31) = 0.34, p = .57, \eta_p^2 = 0.01 [0, 0.17]$ <sup>32</sup>; cue-type:  $F(1, 31) = 0.35, p = .56, \eta_p^2 = 0.01 [0, 0.17]$ ; interaction :  $F(1, 31) = 0.08, p = .78, \eta_p^2 = 0.003 [0, 0.13]$ ).

*Type-2 ROC curves.* Analyses conducted on AUROC2 showed a different pattern of results (see Figure 6.5b). AUROC2 were smaller for pairs with pseudoword targets than for pairs with word targets ( $F(1, 31) = 7.19, p = .01, \eta_p^2 = 0.19 [0.01, 0.42]$ ). AUROC2 were also smaller for pairs with pseudoword cues compared to pairs with word cues ( $F(1, 31) = 11.81, p = .002, \eta_p^2 = 0.28 [0.05, 0.50]$ ). Moreover, we observed an interaction ( $F(1, 31) = 6.68, p = .02, \eta_p^2 = 0.18 [0.01, 0.41]$ ) with the effect of target-type being significant for pairs with a word cue ( $t(31) = 3.65, p < .001, d_z = 0.65 [0.26, 1.04]$ ) but not for pairs with a pseudoword cue ( $t(31) = 0.23, p = .82, d_z = 0.04 [-0.31, 0.39]$ ). As for gamma scores analyses, all of these effects disappeared

<sup>31</sup> For two participants, gamma correlation could not be calculated in one condition (1 in the PW/PW condition and 1 in the W/W condition).

<sup>32</sup> The residuals of the model were not normally distributed. Wilcoxon test :  $V = 316, p = .34$ .

when only considering unrecalled trials (target-type:  $F(1, 31) = 0.06$ ,  $p = .81$ ,  $\eta_p^2 = 0.002$  [0, 0.12]; cue-type:  $F(1, 31) = 3.72$ ,  $p = .06$ ,  $\eta_p^2 = 0.11$  [0, 0.33]; interaction :  $F(1, 31) = 0.59$ ,  $p = .45$ ,  $\eta_p^2 = 0.02$  [0, 0.19]).

*Metacognitive efficiency.* As for the other experiments, MCMC chains in the hierarchical modelling of the estimates of the effects of target-type and cue-type did not converge. Graphs could thus not be interpreted in a reasonable manner. As we preregistered such analyses, we provided the results in [Appendix I](#).

### 6.5.3. Interim discussion

As expected and in line with previous work on associative learning showing beneficial aid from semantic processing on recall performance (Bower & Winzenz, 1970; Straub & Granaas, 1992), participants recalled more targets from pairs containing two meaningful items than any other pairs. In fact, very few participants were able to recall meaningless targets or targets associated with meaningless cues. This can be explained by the fact that in conditions in which at least one item of the pair was meaningless such items cannot rely on semantic elaboration. As such, no semantic properties can be used to create an effective association with the other item of the pair, which could reduce memory strength and therefore lead to reduced recall (e.g., no use of imagery, no sentence generation).

Metacognitive sensitivity mirrored recall performance but only when looking at AUROC2 measure: participants showed better sensitivity in predicting their recognition for pairs with two meaningful items but their sensitivity worsened when at least one item was meaningless. However, this effect disappeared when only looking at trials with no recall. This suggests that the effect of meaninglessness on metacognitive sensitivity comes from the recall attempt. Indeed, it suggests that participants were better at predicting their performance on the recognition test for trials of pairs with two meaningful items because they had access to the target in some cases. This result will be discussed in the General Discussion.

Contrary to Experiment 6.2 and our hypothesis, increasing presentation time of meaningless items did not improve recognition of meaningless targets. However, unrecalled meaningful targets were not significantly more recognized than unrecalled meaningless targets. This means that when participants were unable to recall the words, they displayed the same recognition performance as for unrecalled pseudoword targets. Like in all previous experiments, targets cued by meaningless items were less well recognized than targets cued by meaningful items, even when they were unrecalled. This effect will be discussed in the [General Discussion](#) of this chapter.

In line with previous experiments and work on cue-familiarity (Schwartz & Metcalfe, 1992), participants predicted worse recognition when a meaningless cue was presented than a meaningful cue, even when they were unable to recall the target. They also predicted more recognition of meaningful targets when they saw a meaningful cue. This last result suggests once again that when they know the meaning of the cue, they use the type of target as an information for their PoR. However, this effect was not significant for unrecalled items.

## 6.6. General discussion of Chapter 6

The four experiments described in this paper explored the importance of semantic processing in metacognitive abilities by comparing word and pseudoword cues and targets. We aimed at evaluating whether access to semantic properties of cues and targets in paired associate learning aids the accurate prediction of one's own recognition performance. Of importance, we did not observe the same effects of manipulating the type of material when considering the different measures of metacognition (i.e., magnitude and sensitivity, see Tableau 6.2).

**Tableau 6.2 :** Summary of results from Chapter 6

	First-order performance (recognition)			Magnitude of PoR			Metacognitive sensitivity		
	Target-type	Cue-type	Interaction	Target-type	Cue-type	Interaction	Target-type	Cue-type	Interaction
<b>Exp 6.1</b>	✓	✓	✗	✓	✓	✓	✓	Equivocal	✓
<b>Exp 6.2</b>	Control	✓ ≠	✓	✗	✓	✓	✓	Equivocal*	✓
	TM	✓ ≠	✓	✗	✓	✓	✓	✓*	Equivocal
<b>Exp 6.3</b>	Control	✗	✓	✗	✓	✓	✓	✓	Equivocal
	Fam	✓	✓	✗	✓	✓	✓	✓	Equivocal
<b>Exp 6.4</b>	All trials	✓	✓	✗	✓	✓	✓	Equivocal	Equivocal
	Unrecalled trials	✗	✓	✗	✓	✓	✗	✗	✗

✓ = significant effect; ✗ = not significant effect

Equivocal = depends on the measure used to assess metacognitive sensitivity

≠ = effect in opposite direction

\* = no significant interaction between groups despite different results

### *Magnitude of predictions as a function of type of cue and access to target*

Results from all experiments show that people use the type of cue to make a prediction of future retrieval. Indeed, participants predicted poorer recognition when a meaningless cue was presented. As participants actually did recognize fewer targets which they learnt with meaningless cues, this can be interpreted as making appropriate metacognitive judgements. When the cue was meaningful, participants also made their predictions in accordance to the

type of target: they predicted more recognition of words than of pseudowords. These results are consistent with the hypothesis developed for the FOK (which is a specific type of PoR) that states that both cue-familiarity and accessibility to target influence FOK (Koriat & Levy-Sadot, 2001). Less familiar or novel cues are usually associated with lower predictions (e.g., Metcalfe et al., 1993; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). Familiarity and frequency of words are closely linked (Noble, 1954; Tanaka-Ishii & Terada, 2011). Therefore, it is possible that the effect of cue-type on the magnitude of PoR comes from the fact that participants less encountered pseudowords than words (or not at all in most experiments except the Fam-group of Experiment 6.3) before they learnt the pairs and made their predictions in the PoR phase. Moreover, when participants are confronted with a cue that is familiar enough (i.e., when it was a word here), they can engage in active search of the target or any partial information about it. In this case, participants modified their PoR based on the type of target. As words can be processed semantically, more features can be accessed about them even when not recalled, as in Experiment 6.4, than about pseudowords for which only orthographic or phonological information can be used. However, when they were not able to recall the complete target, participants did not use the type of target as a cue to guide their PoR even if they were still influenced by the type of cue. This cannot allow us to conclude that more partial information about meaningful targets were accessed compared to meaningless targets when they were not recalled.

In sum, both actual recognition performance and PoR are lower when the cue cannot be semantically processed, even when the target cannot be recalled. We suggest this effect comes from the fact that pseudowords are new and less familiar than words. Moreover, when cues are meaningful, no matter the actual first-order performance, PoR are higher for meaningful targets (i.e., W/W condition). However, this specific target-type effect seems to derive from the fact that participants are actually able to retrieve the targets when making their PoR, which make them predict accurate recognition of these targets. Nonetheless, it is possible that explicitly asking participants to recall the target before making their predictions in Experiment 6.4 could have driven them to use the retrieval outcome (i.e., success or failure) as a cue more than not when they were not asked to (Experiments 6.1, 6.2, 6.3). In this case, people could have been more cautious/conservative in making their predictions. Our results also show that people do not need to be explicitly asked to recall the target to be influenced by the same types of information as the FOK, which is a PoR for unrecalled items only (i.e., type of cue and amount of information available about the target). Nonetheless, the fact that even when people fail to recall the target, magnitude is still influenced by the type of cue and the type of target confirm that people still have some awareness of unretrieved information when predicting future retrieval.

*Importance of semantic processing for metacognitive sensitivity: the role of recall*

These experiments show that being able to semantically process both the cue and the target of a pair is important for metacognitive sensitivity. Indeed, participants were less metacognitively sensitive when at least one item of the pair could not be semantically processed. Such an effect seems to come from the fact that targets are less recalled in this case. Previous work shows that semantic processing strengthens memory (e.g., levels of processing, Craik & Lockhart, 1972; Lockhart & Craik, 1990). Efficient associative learning linked to the quality of original encoding is especially important for metacognitive sensitivity of predictions (Hertzog et al., 2010; Lupker et al., 1991; Thomas et al., 2012). For example, studies have shown a deficit in the eFOK in aging (Morson et al., 2015; Perrotin et al., 2006; Souchay et al., 2007) that could result from a decrease in the quality of encoding (Hertzog et al., 2010). The impact of recall attempts on metacognitive sensitivity have been corroborated by previous work showing that metacognitive efficiency is correlated with recall performance in an FOK task (Mazancieux, Dinze, et al., 2020) and that metacognitive sensitivity is better when FOK are not only made for unrecalled but also on for recalled trials (Schwartz et al., 2016). In Experiment 6.4, we explicitly asked participants to recall the target but they might also have accessed it before making their PoR even when not explicitly asked to in the other three experiments. Yet, once again, explicit recall attempts could change the cues used to make PoR and make participants favor actual recall success as a cue, which is highly predictive of future recognition, more than when they are not explicitly asked to retrieve the target. Thus, we cannot know for sure that complete retrieval of the target was the source of changes in metacognitive sensitivity between conditions in Experiments 6.1, 6.2 and 6.3.

Previous research suggests that even when a target is not retrieved some semantic features can be accessed (Koriat et al., 2003). Accessing such semantic properties (whether accurate or not) can in turn influence one's ability to predict one's performance (Thomas et al., 2012). However, metacognitive sensitivity was not influenced by whether the cue or the target could be semantically processed when they were not recalled. Therefore, we cannot conclude that even when they could not recall the target, participants used semantic partial information about the target to make their PoR. Note however that partial recall was not directly measured so we cannot attest that semantic properties were not used for meaningful targets. It is possible that different cues were used for meaningful and meaningless targets. Indeed, the fact that semantic features cannot be used for meaningless targets does not exclude other types of information to be used to infer future recognition (e.g., episodic details, orthographic cues).

Finally, as we used concrete words, we can question which aspect of semantic processing is important for the cue-target association and its influence on recall and PoR (e.g., imagery, sentence generation or another strategy). We could therefore wonder whether the same

results would be obtained using abstract words. Abstract words can be semantically processed but do not involve perceptual features. As shown by the concreteness effect, concrete words are usually better remembered than abstract words (Paivio et al., 1994). It could therefore be interesting to test whether semantic processing is important for PoR beyond perceptual features and visual mental imagery by comparing metacognitive sensitivity for abstract words and meaningless words.

#### *The first-order and second-order relationship*

Our results on first-order performance were not congruent with the previous studies which failed to find difference in performance based on lexicality, even if recognition of pseudowords relied more on familiarity and recognition of words on recollection (Gardiner & Java, 1990; Rajaram et al., 2002; Experiment 2 of Geraci et al., 2009 but see Haaf et al., 2021). However, one previous study also found better recognition of words than pseudowords (Experiment 1 of Geraci et al., 2009). As the global recognition performance for pseudowords and words was not the primary goal of these papers, it is unclear which factors drive such divergence.

As reflected by the performance on the recognition test in all experiments, participants remembered targets cued by meaningful cues better than targets cued by meaningless cues. One hypothesis is that cues were presented on the left part of the pair and as French is read from left to right, cues were therefore read before the targets (especially in Experiment 2 and 4 in which cues were presented first) which were therefore not well processed. We can wonder if participants spent more time on the cues when they were meaningless, therefore not processing the target correctly. The effect of cue-type observed on the magnitude of PoR might therefore be explained by such an effect. Overall, as first-order performance was not equated between conditions and might therefore be confounded in metacognitive performance (Fleming & Lau, 2014; Vuorre & Metcalfe, 2021), our interpretations on second-order performance need to be further explored in paradigms with equivalent first-order performance. Nonetheless, our results show that when people cannot semantically process the information they need to remember, they are less metacognitively sensitive to predict future retrieval and this can be explained by lower memory strength and lower recall, thus memory factors.

## CHAPITRE 6 : CONCLUSION

In four experiments, we showed that participants infer future recognition from whether they have a **semantic representation** of the cue that is presented when making their prediction. Indeed, whether they are able to recall the target or not, they predict less recognition when meaningless cues are presented. Moreover, when they are familiar with the cue, they also use the type of target as a cue to make their predictions, which do not always reflect the actual memory pattern.

Our results also support the claim that participants are more sensitive in predicting the future recognition of newly learnt information when they have a representation of the items stored in memory (i.e., when they can semantically process them). Our results show that no matter how well they perform on the first-order task (i.e., recognition) what matters is that a **semantic association** can be formed between the cue that gives rise to the prediction and the target and they can retrieve the target. However, when the target cannot be recalled, we cannot conclude that semantic processing of neither the cue nor the target influences metacognitive sensitivity. Therefore, we propose that semantic processing helps people to be more metacognitively sensitive as a consequence of a stronger memory trace.

# Chapitre 7.

## Effet de la fluence conceptuelle sur la métacognition épisodique

---

---

Dans le chapitre précédent, nous avons montré que le traitement des informations sémantiques était important pour prédire une future récupération en mémoire épisodique. Ce chapitre 7 vise à évaluer comment les connaissances que les individus possèdent sur le monde et sur les informations à traiter peuvent influencer leur auto-évaluation de l'exactitude de leurs souvenirs. Dans deux études, nous avons cherché à évaluer l'influence de l'utilisation des informations sémantiques sur la métacognition dans des tâches de mémoire épisodique. L'intervention des connaissances sémantiques a été étudiée en manipulant la congruence sémantique d'images.

Dans l'expérience 7.1, l'objectif était de manipuler la congruence sémantique des cibles et des distracteurs au moment du test de mémoire afin de faire varier la fluence conceptuelle de ces items, et ainsi, de voir comment cela affectait les RCJ. Nous faisons l'hypothèse que cette manipulation n'aurait pas d'effet sur la performance de premier ordre. Le but était de voir si la congruence sémantique était tout de même utilisée comme indice par les individus pour évaluer l'exactitude de leurs réponses de type-1, même dans un cas où la performance de premier ordre n'était pas influencée par cette manipulation, et ainsi créer une illusion métacognitive (i.e., modification des jugements métacognitifs sans modification de la performance de premier ordre).

Dans l'expérience 7.2, l'objectif était de manipuler la congruence sémantique des images lors de l'encodage afin de voir comment cela affectait la mémoire de ces items, mais surtout, la capacité des participants à prédire leur performance mnésique. Les participants ont donc dû faire des PoR pour les items appris dans des scènes congruentes ou incongruentes. Nous nous attendions à ce que la manipulation de la fluence conceptuelle à l'encodage influence la mémoire, et nous cherchions à voir si les participants arrivaient à prédire cet effet.

Les résultats de l'Expérience 7.1 ont été présentés au colloque de l'Association for the Scientific Study of Consciousness (ASSC), Amsterdam, Pays-Bas, Juillet 2022 ; et au colloque de l'European Society for Cognitive Psychology (ESCOP), Lille, France, Août 2022.
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 7.1. Introduction

Pour rappel, afin d'évaluer leurs propres performances sur une tâche de mémoire, des individus peuvent utiliser plusieurs indices (Koriat, 1997). De précédentes études suggèrent que la fluence, c'est-à-dire la facilité subjective avec laquelle une information est traitée, peut servir d'indice et influencer les jugements métacognitifs. Différentes sources de fluence peuvent être étudiées comme la fluence perceptuelle, la fluence de récupération, et la fluence conceptuelle (Alter & Oppenheimer, 2009). La fluence conceptuelle dépend du degré de congruence de l'information à traiter avec les connaissances préalables qu'une personne a sur le monde. L'étude de la fluence conceptuelle constitue donc un moyen d'évaluer l'influence de ces connaissances préalables (i.e., la mémoire sémantique) sur les jugements métacognitifs et les capacités métacognitives.

Comme il en a été discuté dans le [Chapitre 3](#), plusieurs études ont mis en évidence l'influence de la fluence conceptuelle (e.g., *relatedness effect*, Castel et al., 2007; Hertzog et al., 2002; Matvey et al., 2006; see Mueller et al., 2013 pour une petite revue) sur la magnitude des jugements d'apprentissage (JOL). Par exemple, dans certaines études, la fluence conceptuelle a été manipulée en faisant varier le degré d'association entre deux mots de paires à apprendre. Une paire de mots non sémantiquement reliés (e.g., banc – fraise) peut être considérée comme non-fluente, c'est-à-dire plus difficile à traiter, par rapport à une autre dont les mots sont reliés (e.g., banc – parc). Pour rappel, ces études ont montré que les participants faisaient des JOL plus élevés pour les paires qui étaient sémantiquement reliées que pour celles qui n'étaient pas sémantiquement reliées (i.e., effet de *relatedness*; Castel et al., 2007; Connor et al., 1997; Dunlosky & Matvey, 2001; Undorf & Erdfelder, 2015). Une autre étude a également montré que la sensibilité des JOL était plus faible pour les paires non-reliées que pour les paires reliées (Hertzog et al., 2002). Ainsi, ces travaux montrent que la fluence conceptuelle peut être utilisée comme indice pour faire des JOL et donc avoir un effet sur leur magnitude, et qu'elle peut également influencer la sensibilité de ces jugements.

Le traitement sémantique des informations à retenir peut avoir un effet sur la performance à la tâche de premier ordre puisqu'il peut influencer la mémoire épisodique (e.g., effet de *relatedness* sur la mémoire, Epstein et al., 1975; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2003, 2005). Dans ce contexte, il est intéressant d'évaluer si les personnes ont connaissance de ces effets, et ainsi de voir si elles arrivent à adapter leurs jugements métacognitifs de manière adéquate. L'expérience 7.2 visait notamment à approfondir cette question. Cependant, la fluence de traitement est un indice qui n'est pas toujours diagnostique de la performance de premier ordre. Cela peut être illustré par un effet de *relatedness* sur la magnitude des JOL également retrouvé lorsque les paires sémantiquement non-reliées étaient en réalité mieux rappelées (Carroll et al., 1997). La manipulation de la fluence conceptuelle

peut donc influencer la magnitude des jugements métacognitifs même dans les cas où la performance réelle n'est pas modifiée par cette manipulation. Ainsi, ces résultats montrent que les connaissances sémantiques sur le matériel à encoder en mémoire épisodique peuvent parfois amener des personnes à faire des erreurs dans l'évaluation de leur performance et créer des illusions métacognitives. Dans l'expérience 7.1, le but était de voir si de telles illusions pouvaient être créées par la fluence conceptuelle d'images.

La majorité des études conduites pour évaluer l'effet de la fluence conceptuelle sur les jugements métacognitifs ont porté sur les JOL immédiats. Peu d'études se sont focalisées sur les prédictions faites au moment de la tentative de récupération en mémoire ou sur les jugements rétrospectifs. Une étude a montré que l'effet de *relatedness* sur la magnitude des jugements métacognitifs semble moindre dans un paradigme de FOK que dans le JOL immédiat (Carroll et al., 1997, voir p50). Dans cette étude, où des paires de mots sémantiquement reliés ou non-reliés étaient apprises, le degré d'association était contrasté au degré d'apprentissage. Les paires non-reliées étaient apprises jusqu'à 8 rappels réussis et les paires reliées jusqu'à 2 rappels réussis. Les participants faisaient ensuite une prédiction de rappel (JOL différé). Quelques semaines après, les participants réalisaient ensuite une tâche classique de FOK. Alors que les JOL étaient plus élevés pour les paires reliées, la magnitude des FOK étaient influencée par le degré d'apprentissage et non par le degré d'association entre les mots : elle était plus élevée pour les paires non-reliées mieux apprises que pour les paires reliées moins bien apprises. Ce résultat montre que le FOK est moins influencé par la fluence conceptuelle que le JOL, et plus par des facteurs mnésiques comme le degré d'apprentissage.

Dans une autre étude menée au sein de notre laboratoire (Expérience 3, Chapitre 6 de Mazancieux, 2020), l'effet de la fluence conceptuelle a été évaluée sur des jugements prédictifs faits juste avant la décision de premier ordre (i.e., PoR) et sur des RCJ. Dans cette étude, les participants devaient mémoriser des patterns visuels insérés dans des mots (voir Figure 3.5 p.52 pour un exemple). Lorsque les mots étaient représentés, ils devaient prédire leur capacité à reconnaître le pattern visuel (uniquement dans la condition PoR). Ensuite, ils devaient identifier le bon pattern visuel parmi deux propositions. Après avoir répondu, les participants devaient indiquer à quel point ils étaient certains que leur réponse était correcte (uniquement dans la condition RCJ). Dans cette étude, la fluence conceptuelle était manipulée par la lexicalité des mots dans lesquels les patterns étaient présentés. Ces mots pouvaient être des pseudo-mots (e.g., « lyk ») ou des mots existants (e.g., « one »). Alors que la magnitude des PoR n'était pas influencée par la lexicalité, les RCJ étaient plus élevés pour les mots que pour les pseudo-mots. De plus, l'efficacité métacognitive était meilleure pour les mots que pour les pseudo-mots dans le cas où les jugements étaient rétrospectifs mais la

lexicalité n'avait d'effet sur l'efficacité métacognitive pour les PoR. Les résultats de cette étude montrent que la fluence conceptuelle peut influencer la magnitude des jugements rétrospectifs mais également l'efficacité métacognitive, en particulier des RCJ.

Dans les études ayant porté sur la fluence conceptuelle dans les paradigmes de métacognition épisodique, la fluence conceptuelle était manipulée à l'aide de matériel verbal (e.g., Mazancieux, 2020 ; voir Mueller et al., 2013 pour les études sur l'effet de *relatedness*). La mémoire sémantique a longtemps été mise en lien avec la compréhension du langage et comme nous l'avons montré dans le [Chapitre 6](#) précédent, le traitement sémantique des informations à retenir est important pour avoir de bonnes performances métacognitives sur des tâches réalisées avec du matériel verbal. Cependant, selon les modèles récents de la mémoire sémantique, la capacité à attribuer un sens aux éléments du monde ne repose pas nécessairement sur le langage (e.g., Binder & Desai, 2011; Lambon Ralph et al., 2017; Martin, 2007, 2016). Notamment, des études menées sur la congruence sémantique (i.e., adéquation entre les stimuli à traiter et les connaissances sémantiques) ont mis en évidence l'influence des connaissances sémantiques sur le traitement visuel d'images (Bonitz & Gordon, 2008; Hollingworth & Henderson, 2000; LaPointe et al., 2013; LaPointe & Milliken, 2016; Loftus & Mackworth, 1978; Ortiz-Tudela et al., 2017). Dans certains de ces paradigmes, les participants devaient indiquer lorsqu'ils détectaient l'apparition d'un item dans une scène. Ces scènes étaient congruentes (e.g., un pré) ou incongruentes (e.g., une route) avec les cibles (e.g., une vache). La détection des cibles étaient plus rapides dans les essais incongruents (Hollingworth & Henderson, 2000; LaPointe et al., 2013; LaPointe & Milliken, 2016) mais leur identification (i.e., indiquer de quel objet il s'agit) plus correcte dans les essais congruents (LaPointe et al., 2013; Ortiz-Tudela et al., 2017, mais voir LaPointe & Milliken, 2016 pour une absence d'effet significatif). Cela suggère donc que les connaissances sémantiques sur le contexte dans lequel des objets/animaux sont généralement rencontrés dans la vie réelle influencent la manière dont ils sont traités. D'autres études mettent également en avant l'influence des représentations sémantiques dans le traitement rapide de scènes visuelles complexes (e.g., Greene et al., 2015). Ainsi, ces résultats montrent que la congruence sémantique d'images joue sur la facilité à traiter un stimulus. Cela suggère donc que la variation de la congruence sémantique d'images peut être utilisée comme manière de manipuler la fluence conceptuelle.

Les capacités métacognitives ont peu été évaluées sur des tâches de mémoire portant sur des images. Une précédente étude a montré que, dans une tâche de mémoire de scènes naturelles, les JOL étaient précis puisque les participants arrivaient bien à prédire leur performance (Undorf & Bröder, 2021). Cependant, à notre connaissance, aucune étude n'a actuellement évalué l'effet de la fluence conceptuelle sur les jugements métacognitifs avec l'aide de matériel imagé ayant du sens (e.g., objets réels). Dans les expériences de ce

chapitre, la fluence conceptuelle a été manipulée dans deux études par la congruence sémantique d'images d'items insérés dans des scènes. L'objectif de ces deux études était d'évaluer l'effet de la fluence conceptuelle sur les jugements métacognitifs rétrospectifs (Expérience 7.1) et les jugements prédictifs faits au moment de la récupération en mémoire (Expérience 7.2), sur des tâches de mémoire épisodique d'images.

Dans l'expérience 7.1, l'objectif de était de voir si la manipulation de la congruence sémantique des cibles et des distracteurs au moment du test de mémoire influençait la confiance que les participants avaient en leurs réponses, dans un contexte où leur performance mnésique n'était pas censée être affectée par cette manipulation. Dans cette étude, les participants mémorisaient les images d'objets présentés seuls. Dans une seconde phase, chaque cible était présentée avec un distracteur. Ils devaient sélectionner la cible et faire un RCJ. Les deux items étaient présentés dans des scènes sémantiquement congruentes (i.e., fluentes) ou incongruentes (i.e., non-fluentes). Nous faisons l'hypothèse que la congruence des scènes des cibles et des distracteurs affecterait la fluence de traitement de ces items sans pour autant impacter la performance de premier ordre. Les items ayant été appris sans contexte à l'encodage, nous ne pensions pas que la manipulation de la congruence au moment du test jouerait sur des processus d'encodage, de rétention, ni de récupération. Ainsi, nous ne nous attendions pas à observer un effet de la congruence sémantique sur le taux de reconnaissance des items. D'un autre côté, nous faisons l'hypothèse que la manipulation de la congruence sémantique jouerait sur le traitement des cibles et des distracteurs, et que la congruence relative des deux images servirait d'indice pour les RCJ. Dans une étude précédente (Mazancieux, 2020), l'efficacité métacognitive des RCJ était plus élevée pour les essais congruents (i.e., patterns visuels dans des mots) que pour les essais incongruents (i.e., patterns visuels dans pseudo-mots). Sur la base de ces résultats, nous prédisions que les participants auraient plus confiance en l'exactitude de leur réponse (i.e., magnitude des RCJ plus élevée) et auraient une meilleure sensibilité métacognitive dans les essais où la cible était plus fluente que son distracteur par rapport aux essais où le distracteur était plus fluent que la cible.

Si les travaux cités précédemment montrent des effets à court terme (i.e., sur la perception) de la congruence sémantique d'images, celle-ci a également des effets à plus long terme, notamment sur la mémoire de ces images (Ortiz-Tudela et al., 2017). En effet, les items présentés dans des contextes congruents étaient mieux retenus que des items présentés dans des contextes incongruents (Ortiz-Tudela et al., 2017). À notre connaissance, aucune étude n'a évalué si les individus ont connaissance de cet effet de congruence sémantique lorsqu'ils réalisent une tâche de mémoire et s'ils peuvent adapter leurs jugements en adéquation avec cet effet. Le but de l'expérience 2 était donc de manipuler la fluence

conceptuelle au moment de l'apprentissage des images d'items et voir l'effet de cette manipulation sur leur mémoire et plus particulièrement sur la capacité des participants à prédire cet effet à travers des PoR. Dans cette étude, les participants devaient mémoriser des items présentés dans des scènes congruentes ou incongruentes. En revoyant chaque scène, ils prédisaient ensuite leur capacité à pouvoir se rappeler de l'item qui était dedans. Sur la bases des résultats de deux études précédentes (Expériences 2 et 3, Ortiz-Tudela et al., 2017), nous faisons l'hypothèse que les participants reconnaîtraient moins d'items congruents que d'items incongruents. Comme discuté dans le [Chapitre 2](#), les PoR peuvent être influencées par la familiarité subjective ressentie vis-à-vis des indices présentés pour faire la prédiction (i.e., scène ici). Si la familiarité avec l'indice est suffisante, les participants peuvent s'engager dans le recherche de la cible et tenter de récupérer des informations sur celle-ci, qui peuvent influencer la magnitude des PoR (Koriat & Levy-Sadot, 2001). De précédentes études (Expérience 2 et 3, Ortiz-Tudela et al., 2017) ont montré que la mémoire des scènes n'était pas influencée par leur congruence avec les items. Ainsi, nous prédisions que si les participants se basaient principalement sur la familiarité avec l'indice (i.e., la scène), aucun effet de congruence ne devrait être observé sur la magnitude des PoR. En revanche, si la récupération (même partielle) de la cible était également prise en compte, les PoR pourraient être plus élevés dans la condition congruente que dans la condition incongruente.

**Tableau 7.1** : Récapitulatif des deux expériences.

	Type de jugement	Manipulation de la fluence	Hypothèses
<b>Expérience 7.1</b>	RCJ	Au moment du test : cible et distracteur	<ul style="list-style-type: none"> <li>• RCJ plus élevés quand la cible est traitée de manière plus fluente que son distracteur et plus faibles lorsque le distracteur est plus fluent que la cible</li> <li>• Pas d'effet de la fluence sur la performance de premier ordre</li> </ul>
<b>Expérience 7.2</b>	PoR	Dès l'encodage	<ul style="list-style-type: none"> <li>• PoR plus élevées pour les essais congruents</li> <li>• Meilleure mémorisation des items congruents</li> </ul>

## 7.2. **Expérience 7.1 : Effet de congruence sémantique sur les jugements rétrospectifs de confiance épisodique**

Le but de l'expérience 7.1 était d'évaluer l'utilisation de la congruence sémantique comme indice dans les RCJ, dans le cas où cette congruence n'affecte pas nécessairement la performance de premier ordre. Nous faisons l'hypothèse que la magnitude, le biais et la sensibilité des RCJ seraient particulièrement affectés dans les cas où la cible et son distracteur n'étaient pas présentés dans une scène du même degré de congruence. Nous attendions à ce que les participants soient particulièrement perturbés pour faire leur RCJ (plus faible magnitude et moins bonne sensibilité) dans les essais avec la cible incongruente et son distracteur congruent (i.e., cible peu fluente rapport au distracteur). À l'inverse, nous faisons l'hypothèse qu'ils seraient meilleurs pour faire leur RCJ (magnitude plus élevée, meilleure sensibilité) dans les essais où la cible était plus fluente que le distracteur (i.e., cible congruente et distracteur incongruent).

### 7.2.1. **Méthode**

La méthode de cette étude a été pré-enregistrée sur l'Open Science Framework (<https://osf.io/9gm8t/>).

#### **Participants**

Cent-trois participants de langue maternelle française âgés de plus de 18 ans ( $M_{age} = 26.36$ ,  $ET_{age} = 5.37$ , 56 femmes) ont été recrutés sur la plateforme Prolific (<https://www.prolific.co/>). Tous les participants ont rapporté ne pas avoir de trouble mental. Le nombre de participants à recruter a été déterminé en réalisant une analyse de puissance sur G\*power pour observer un effet de taille  $d = 0.28$  avec une puissance de .80 et un seuil  $\alpha = .05$ . Cette taille d'effet était celle de la fluence conceptuelle sur la magnitude observée dans une précédente étude de notre laboratoire (thèse de Audrey Mazancieux, 2020, Expérience 3, Chapitre 6). Cependant, les données de 26 participants ont été exclues des analyses (voir section [Résultats](#) pour le détail). Ainsi, les données de 77 participants ( $M_{age} = 26.25$ ,  $ET_{age} = 5.55$ , 40 femmes) ont été incluses dans les analyses.

#### **Matériel**

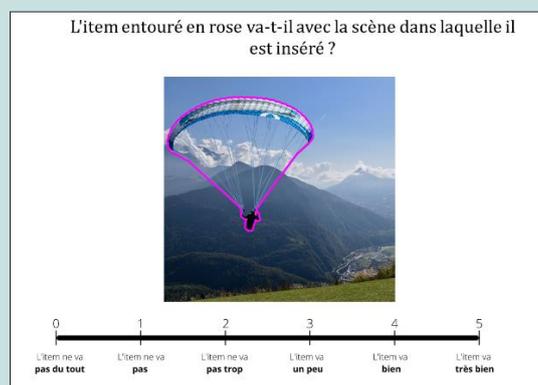
Les photos de 360 items (e.g., animaux, aliments, objets) ont été sélectionnées pour les besoins de cette étude (e.g., un bateau). Pour chaque item, deux autres photos ont été sélectionnées : une scène sémantiquement congruente (e.g., la mer) et une scène sémantiquement incongruente (e.g., un champ). Toutes les scènes ont été égalisées en luminance et contraste sur MATLAB (The Math Works, Inc., 2020). Chaque item a ensuite été

photoshopé dans ses deux scènes respectives (i.e., congruente et incongruente), à la même place sur l'image. Pour que les items soient repérés plus facilement dans l'image, ils étaient entourés d'une ligne rose. Au total, nous avons donc créé 720 images (360 congruentes et 360 incongruentes). Sur la base d'une étude pilote (voir **Encart 2.** ), nous avons sélectionné 336 items et leurs deux montages respectifs (672 images au total). Certaines des images non-sélectionnées ont été modifiées pour être plus (in)congruentes afin d'être utilisées pour la phase d'entraînement du début de l'expérience.

### Encart 2. Étude pilote pour la sélection des stimuli de l'expérience 7.1

Plusieurs bases d'images d'objets insérés dans des contextes sémantiquement congruents et incongruents sont disponibles en ligne (e.g., SCEGRAM, Öhlschläger & Vö, 2017 ; ScenObjects, Roux-Sibilon & Peyrin, 2019). Cependant, ces bases d'images ne possédaient pas assez d'images par rapport au nombre que nous visions. De plus, puisqu'elles possédaient des critères spécifiques (e.g., objets manipulables ou contrôles de propriétés physiques très spécifiques (e.g., couleurs, nombre de pixels) non nécessaires pour notre étude et complexes à contrôler, nous avons été contraints de construire une nouvelle base d'images pour les besoins des expériences 7.1 et 7.2. Afin de vérifier que les images construites étaient bien perçues comme congruentes ou incongruentes, nous avons réalisé plusieurs études pilotes (voir Box 3 pour les études pilotes pour l'expérience 7.2).

Dans une étude pilote réalisée en ligne, 379 participants français âgés de 18 à 44 ans ont été recrutés ( $M_{age} = 26.16$ ,  $ET_{age} = 5.82$ , 184 femmes). Ils devaient juger la pertinence d'images (voir image ci-dessous). Chaque image a été jugée par 124 à 129 participants sur une échelle de 0 (l'item ne va pas du tout dans la scène) à 5 (l'item va très bien dans la scène). Ce pré-test nous a permis de nous assurer que les images étaient bien considérées comme congruentes ou incongruentes.



Pour identifier les images n'ayant pas été perçues comme ce qui était attendu, nous avons calculé la moyenne des jugements et la proportion de jugements attendus (e.g., 1,2, ou 3 pour les incongruentes ; 3, 4 ou 5 pour les congruentes) pour chacune des images. Puisque certaines images n'étaient pas jugées aussi clairement qu'attendu, nous avons décidé de ne pas les sélectionner pour l'étude finale. En effet, nous avons retiré les images congruentes ayant reçu en moyenne un jugement inférieur à 3 (e.g., épée dans un château) et les images incongruentes ayant reçu en moyenne un jugement supérieur à 2. Un t-test a révélé que les images incongruentes étaient bien perçues comme plus incongruentes ( $M = 0.30$ ,  $ET = 0.29$ ) que les images congruentes ( $M = 4.43$ ,  $ET = 0.48$ ) ( $t(359) = 140.82$ ,  $p < .001$ ).

### Protocole

La tâche a été créée sur la plateforme Gorilla et était décomposée en trois phases (voir Figure 7.1). Dans une première phase les participants devaient mémoriser 168 images d'items seuls (animaux, aliments, objets) présentées pendant 1s. Dans la phase de test, les participants voyaient chacune des cibles (i.e., items appris) insérée dans une scène visuelle avec un nouvel item (i.e., distracteur) inséré dans une autre scène visuelle. Pour faire varier le degré de fluence conceptuelle, les scènes étaient soit sémantiquement congruentes (i.e., fluentes) ou incongruentes (i.e., non-fluents) avec l'item qu'elles contenaient. Les participants devaient sélectionner la scène contenant la cible. Après chaque réponse, ils devaient indiquer à quel point ils étaient certains d'avoir donné la bonne réponse (i.e., RCJ). Pour éviter des performances trop élevées et ajouter du temps entre l'apprentissage et le test, les participants ont réalisé une tâche d'attention visuelle. Dans cette tâche distractive, les participants devaient détecter des ronds présentés à gauche ou à droite de l'écran qui étaient précédés par une flèche pointant vers la gauche ou la droite (i.e., similaire au paradigme de Posner, 1980). Ils devaient indiquer le côté duquel le rond était apparu.

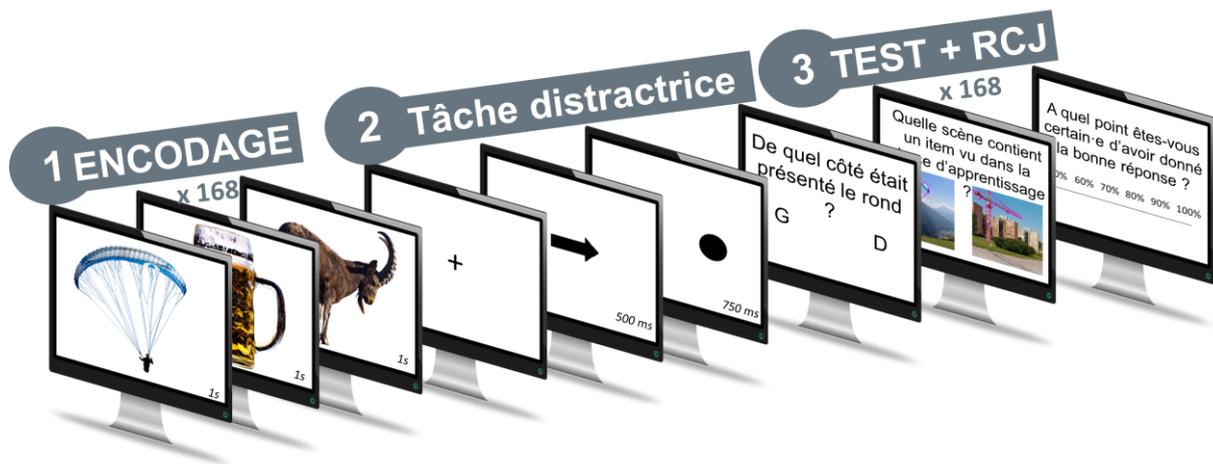
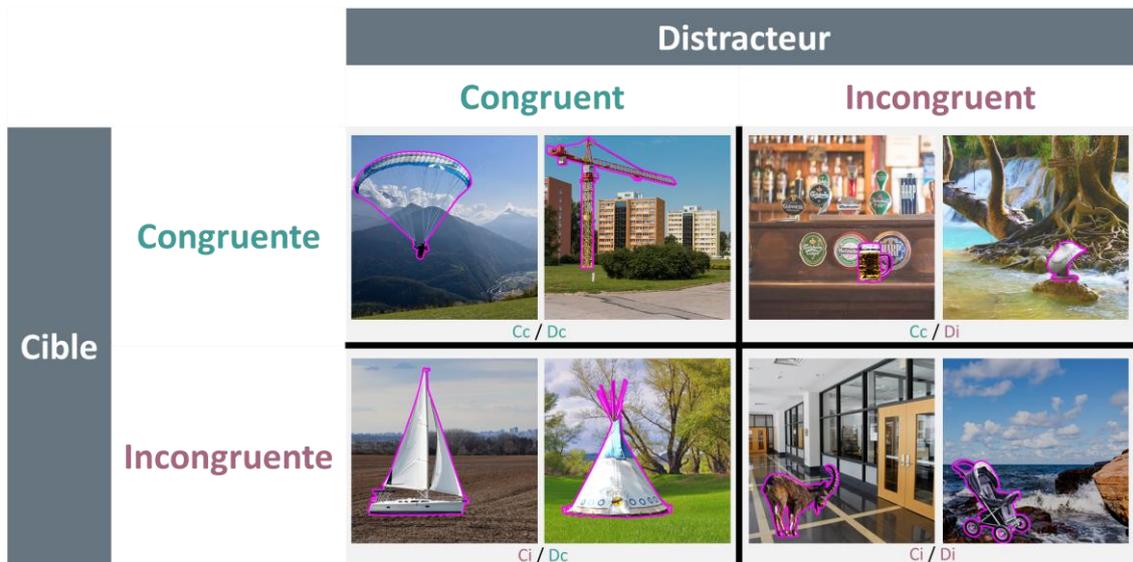


Figure 7.1 : Procédure utilisée dans l'Expérience 7.1.

Dans la phase de reconnaissance, 4 types d'essais étaient présentés dans un ordre aléatoire : a) la cible et le distracteur insérés dans leurs scènes sémantiquement congruentes respectives (Cc/Dc), b) la cible congruente et le distracteur incongruent (Cc/Di), c) la cible incongruente et le distracteur congruent (Ci/Dc), d) la cible et le distracteur incongruents (Ci/Di). Les participants ont réalisé 42 essais de chacune de ces conditions. Afin de limiter les possibles effets de fréquence ou d'effets visuels propres aux items, les cibles et les distracteurs ainsi que la congruence de leurs scènes étaient contrebalancés. Ainsi, les items qui étaient pour

certaines présentées en cibles étaient présentées en tant que distracteurs pour d'autres et inversement. De même, un item (cible ou distracteur) présenté dans une scène congruente pour certains était présenté dans une scène incongruente pour d'autres. Les paires cible-distracteur étaient les mêmes pour tous les participants.



**Figure 7.2 :** Design et exemples d'images utilisées dans l'expérience 7.1. Les cibles sont présentées à droite dans chaque condition du tableau, et les distracteurs à gauche.

### Analyses

Nous avons réalisé des ANOVA pour mesures répétées (congruence de la cible : congruente, incongruente) x 2 (congruence du distracteur : congruent, incongruent). Toutes les analyses ont été effectuées sur le logiciel Rstudio (version 1.3.1073) sur les mesures de proportion de reconnaissance correcte (%RC), la magnitude des RCJ, le biais métacognitif, et la sensibilité métacognitive. Le biais métacognitif était calculé en soustrayant la performance réelle aux jugements moyens. Une valeur négative indique une sous-estimation et une valeur positive indique une surestimation. Le seuil de décision statistique était fixé à  $\alpha = .05$ . Les analyses respectaient les conditions d'application pour l'utilisation de modèles linéaires. Pour vérifier l'absence d'effet significatif, des analyses bayésiennes complémentaires (non pré-enregistrées) ont été réalisées à l'aide du logiciel JASP (JASP team, 2020). Comme pré-enregistré, nous avons également effectué des analyses pour estimer les effets de congruence de la cible et du distracteur ainsi que leur interaction sur l'efficacité métacognitive. Ces effets ont été estimés par des modèles hiérarchiques bayésiens issus des travaux sur le Hmeta-d (Fleming, 2017, <https://github.com/metacoglab/HMeta-d>). La fonction « [Fit\\_metad\\_2wayANOVA.txt](#) » a été utilisée.

### 7.2.2. Résultats

#### **Exclusion de données**

Sur la base des critères pré-enregistrés, nous avons exclu 26 participants des analyses. Trois ont obtenu une performance en mémoire générale au niveau du hasard (inférieur à 55%CR) suggérant un faible effort fourni dans la tâche. Les données de 19 personnes ont également été exclues des analyses en raison de performances plafonds (supérieur à 95%CR) sur la tâche de reconnaissance dans au moins une condition (5 dans la condition Cc/Dc; 9 dans la condition Cc/Di ; 11 dans la condition Ci/Dc ; 8 dans la condition Ci/Di), empêchant le calcul des scores de sensibilité métacognitive. Nous avons également exclu des analyses les données de 4 participants supplémentaires car ils présentaient des données extrêmes (supérieures à 3 ET de la moyenne) sur au moins une mesure étudiée. Par soucis de cohérence, ces données ont été exclus des analyses sur toutes les variables.

#### **Performance en mémoire.**

Les performances de reconnaissance n'ont pas été significativement influencées par la congruence de la cible ( $F(1,76) = 0.21, p = .65, \eta_p^2 = 0.003$ ) ni par la congruence du distracteur ( $F(1,76) = 1.35, p = .25, \eta_p^2 = 0.02$ ), voir Figure 7.3a. L'interaction entre les effets de congruence de la cible et du distracteur n'était pas significative ( $F(1,76) = 1.72, p = .19, \eta_p^2 = 0.02$ ). Une preuve supplémentaire de ces effets nuls a été apportée par une analyse bayésienne complémentaire (non pré-enregistrée ; Rouder et al., 2009). Une ANOVA pour mesures répétées bayésienne a fourni une preuve modérée en faveur de l'hypothèse nulle pour tous ces effets (congruence de la cible :  $BF_{01} = 7.36$ , congruence du distracteur :  $BF_{01} = 4.00$ , interaction :  $BF_{01} = 27.76$ )<sup>33</sup>. Le détail des moyennes est donné en [Annexe J](#).

#### **Magnitude des RCJ**

La magnitude des RCJ était influencée par l'effet de congruence du distracteur. Les RCJ étaient plus élevés lorsque le distracteur était congruent ( $M = 0.774, SD = 0.08$ ) qu'incongruent ( $M = 0.767, SD = 0.08$ ), ( $F(1,76) = 6.07, p = .02, \eta_p^2 = 0.07$ ). Cependant, les effets de la congruence de la cible ( $F(1,76) = 0.34, p = .56, \eta_p^2 = 0.004$ ) et l'interaction entre congruence de la cible et congruence du distracteur ( $F(1,76) = 0.97, p = .33, \eta_p^2 = 0.01$ ) sur la magnitude des RCJ n'étaient pas significatifs (voir Figure 7.3b pour proportions). Des analyses complémentaires (non pré-enregistrées) ont montré le même pattern de résultats lorsque seules les réponses correctes étaient considérées : la magnitude des RCJ était plus élevée pour les essais réussis avec un distracteur congruent ( $M = 0.82, SD = 0.08$ ) que pour les

<sup>33</sup> voir Jarosz & Wiley (2014) pour un tableau récapitulatif pour l'interprétation des facteurs de Bayes, sur la base de Jeffreys (1998)

essais réussis avec un distracteur incongruent ( $M = 0.80$ ,  $SD = 0.08$ ), ( $F(1,76) = 11.38$ ,  $p = .001$ ,  $\eta_p^2 = 0.13$ ).

### **Biais métacognitif**

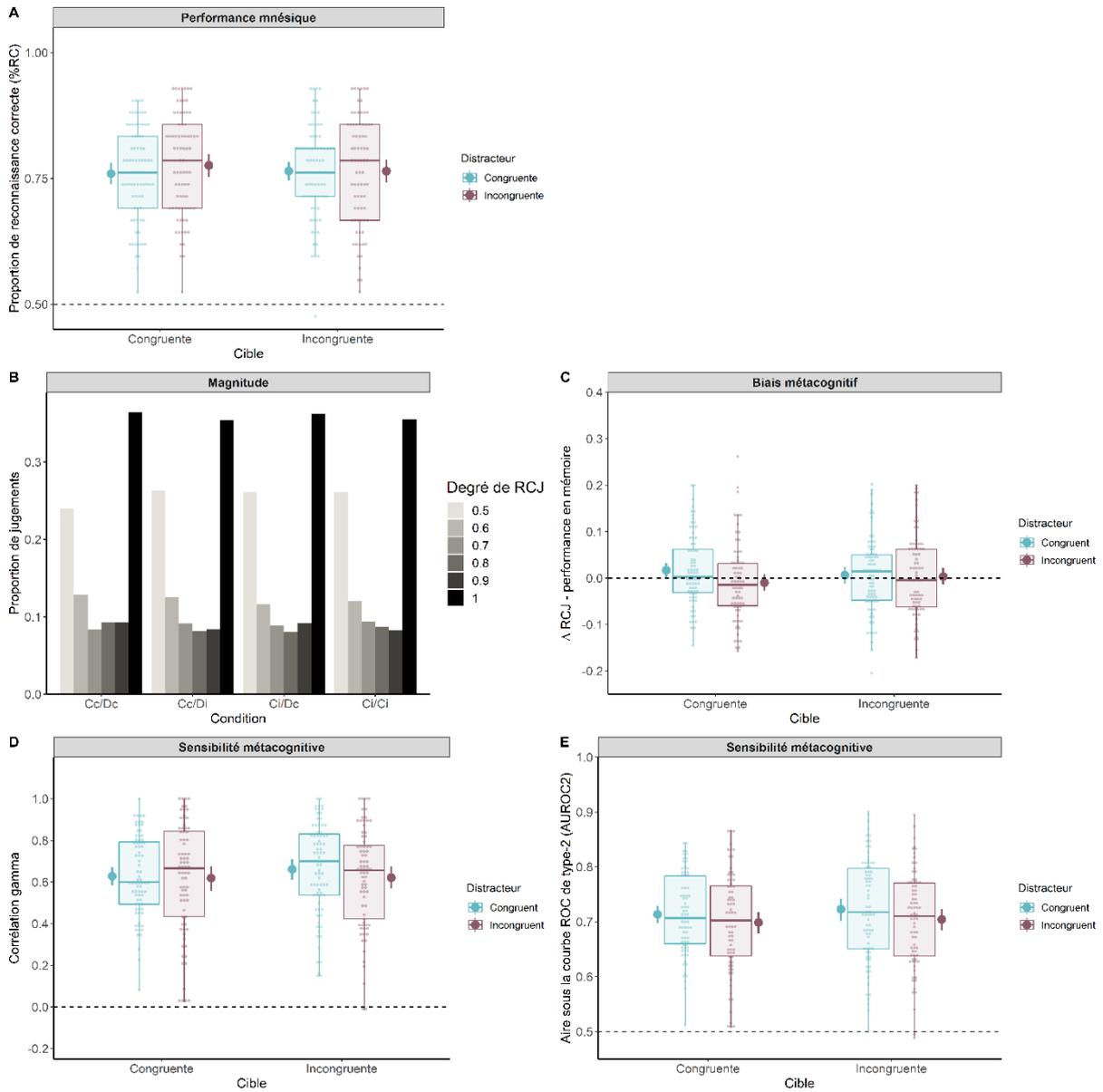
En prenant en compte la performance réelle à la tâche (i.e., biais métacognitif), l'effet du distracteur était toujours significatif mais faible ( $F(1,76) = 4.50$ ,  $p = .04$ ,  $\eta_p^2 = 0.06$ ) puisque les participants étaient légèrement plus enclins à se surestimer lorsque le distracteur était congruent ( $M = 0.01$ ,  $SD = 0.08$ ) que lorsqu'il était incongruent ( $M = -0.003$ ,  $SD = 0.09$ , voir Figure 7.3c). La congruence de la cible n'a pas eu d'effet significatif sur le biais métacognitif ( $F(1,76) = 0.04$ ,  $p = .84$ ,  $\eta_p^2 < 0.001$ ). Cependant, l'effet de congruence du distracteur interagissait avec celui de la cible ( $F(1,76) = 4.47$ ,  $p = .04$ ,  $\eta_p^2 = 0.06$ ). En effet, lorsque la cible était congruente les participants avaient tendance à plus surestimer leur performance quand le distracteur était également congruent ( $M = 0.02$ ,  $SD = 0.08$ ) que lorsque le distracteur était incongruent ( $M = -0.01$ ,  $SD = 0.08$ ), ( $t(76) = 3.20$ ,  $p = .002$ ,  $d_z = 0.37$ ). Cette différence n'était pas significative lorsque la cible était dans un contexte incongruent ( $t(76) = 0.35$ ,  $p = .73$ ,  $d_z = 0.04$ ). Le biais métacognitif n'était significativement différent de zéro dans aucune condition (Cc/Ci :  $t(76) = -1.05$ ,  $p = .30$  ; Ci/Dc :  $M = 0.006$ ,  $SD = 0.08$ ,  $t(76) = 0.68$ ,  $p = .50$  ; Ci/Di :  $M = 0.003$ ,  $SD = 0.09$ ,  $t(76) = 0.31$ ,  $p = .76$ ) mais était proche de la significativité dans la condition Cc/Dc ( $t(76) = 1.94$ ,  $p = .055$ ,  $d_z = 0.22$ ) dans laquelle les participants avaient tendance à se surestimer.

### **Sensibilité métacognitive**

*Corrélations gamma.* Les scores gammas n'ont pas été significativement influencés par la congruence de la cible ( $F(1,76) = 0.92$ ,  $p = .34$ ,  $\eta_p^2 = 0.01$ ), ni celle du distracteur ( $F(1,76) = 1.86$ ,  $p = .18$ ,  $\eta_p^2 = 0.02$ ). L'interaction entre les deux effets n'était pas significative non plus ( $F(1,76) = 0.46$ ,  $p = .50$ ,  $\eta_p^2 = 0.006$ ), voir Figure 7.3d.

*Aire sous la courbe ROC de type-2.* Un pattern de résultats différent était observé en utilisant l'aire sous la courbe ROC de type-2 comme mesure de sensibilité métacognitive (voir Figure 7.3e). En effet, la congruence du distracteur a eu un effet significatif sur cette mesure ( $F(1,76) = 5.71$ ,  $p = .02$ ,  $\eta_p^2 = 0.07$ ). L'AUROC2 était plus élevée pour les essais dans lesquels le distracteur était congruent ( $M = 0.72$ ,  $SD = 0.08$ ) par rapport aux essais dans lesquels le distracteur était incongruent ( $M = 0.70$ ,  $SD = 0.09$ ). L'effet de la congruence du distracteur sur les AUROC2 a également été testé avec des analyses bayésiennes complémentaires qui n'ont pas permis d'accepter H1 ni l'hypothèse nulle ( $BF_{01} = 1$ ). Comme les scores gammas, les scores AUROC2 n'ont pas été significativement influencés par la congruence de la cible

( $F(1,76) = 1.27, p = .26, \eta_p^2 = 0.02$ ) et l'interaction entre les deux effets de congruence de la cible et du distracteur n'était pas significative ( $F(1,76) = 0.04, p = .83, \eta_p^2 < 0.001$ ).

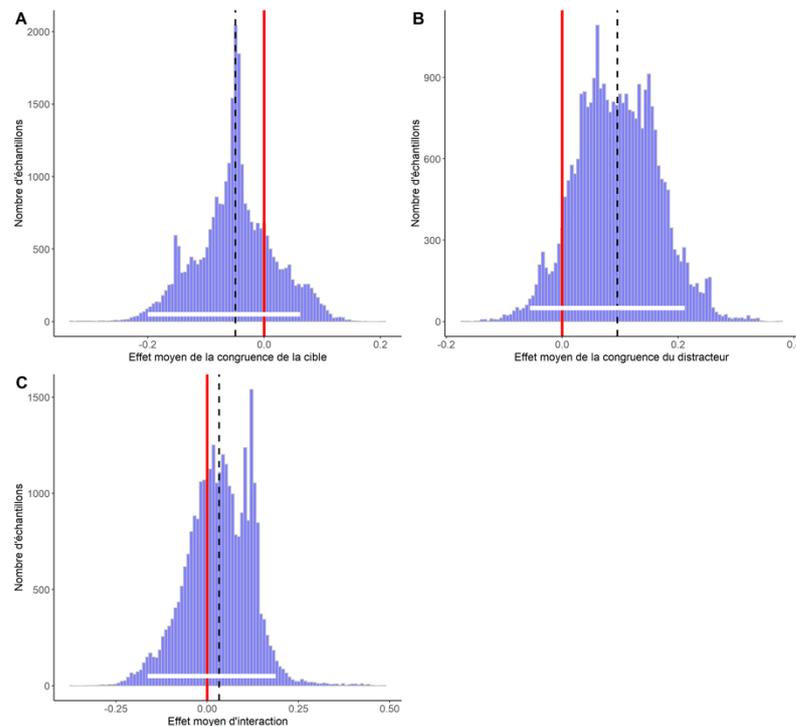


**Figure 7.3 : Résultats de l'expérience 7.1.**

A) Performance de premier ordre, B) Magnitude des RCJ, C) Biais métacognitif, D) Scores gamma, E) AUROC2. Les points épais sont les moyennes par condition, avec l'intervalle de confiance à 95%. Les points plus petits sont les données pour chaque participant.

### Efficiences métacognitive

Les résultats des modèles hiérarchiques bayésiens pour estimer les différents effets sur le M-ratio montre que les intervalles de crédibilité des distributions postérieures (highest density interval, HDI ; le plus petit intervalle contenant 95% des échantillons MCMC ; Kruschke, 2015 ; équivalent des intervalles de confiance en statistiques fréquentistes) se chevauchent avec zéro (voir Figure 7.4). Cela suggère qu'il est peu probable que les effets de la congruence de la cible, du distracteur et leur interaction sur le M-ratio soient différents de zéro.



**Figure 7.4 :** Distributions postérieures des effets moyens ( $\mu$ , ligne en pointillés) de la congruence de la cible (A), de la congruence du distracteur (B), et de leur interaction (C). Les lignes blanches représentent les intervalles à haute densité (95% highest density intervals; HDI). Un recouvrement avec zéro (ligne rouge) suggère une absence d'effet.

### Analyses complémentaires (non pré-enregistrées)

Les analyses sur les temps de réponse ( $TR$ )<sup>34</sup> à la tâche de reconnaissance ont révélé un effet tendanciel de la congruence du distracteur ( $F(1,74) = 3.86, p = .053, \eta_p^2 = 0.05$ )<sup>35</sup>, mais pas de la congruence de la cible ( $F(1,74) = 2.39, p = .13, \eta_p^2 = 0.03$ ), ni d'interaction entre les deux ( $F(1,74) = 2.10, p = .15, \eta_p^2 = 0.28$ ). Les participants étaient plus rapides pour répondre lorsque les distracteurs étaient congruents ( $M_{TR} = 2409\text{ms}, SD_{TR} = 602$ ) que lorsqu'ils étaient incongruents ( $M_{TR} = 2450\text{ms}, SD_{TR} = 623$ ). Le même pattern de résultats était observé lorsque seules les essais corrects étaient considérés (effet de la congruence du distracteur :  $F(1,74)$

<sup>34</sup> Les analyses ont été réalisées sur le logarithme des temps de réponse car les données brutes n'étaient pas normalement distribuées.

<sup>35</sup> Deux participants ont été retirés des analyses en raison de temps de réponse trop lents (> à 3 écarts-types).

= 6.33,  $p = .01$ ,  $\eta_p^2 = 0.08$ ), les effets de congruence de la cible et d'interaction n'étaient pas significatifs ( $F < 1$ ). Les RCJ étaient négativement corrélés aux TR ( $r = -0.21$ ,  $p < .001$ ), dans toutes les conditions (Cc/Dc :  $r = -0.17$ ,  $p < .001$  ; Cc/Di :  $r = -0.22$ ,  $p < .001$  ; Ci/Dc :  $r = -0.21$ ,  $p < .001$  ; Ci/Di :  $r = -0.25$ ,  $p < .001$ ). Les participants avaient donc tendance à donner des RCJ plus élevés pour les essais répondus rapidement.

Afin de déterminer si l'effet de la congruence du distracteur sur la magnitude des jugements pouvaient être expliqués par les différences de TR, une ANCOVA à deux facteurs (congruence de la cible et congruence du distracteur) avec le TR comme covariable a été conduite. Après ajustement par les TR, l'effet du distracteur sur la magnitude des jugements n'était plus significatif ( $F(1,295) = 0.16$ ,  $p = .50$ ). Cela était également le cas lorsque seuls les essais corrects étaient considérés ( $F(1,295) = 1.35$ ,  $p = .25$ ).

### 7.2.3. Discussion de l'expérience 7.1

Le but de cette expérience était de voir comment les connaissances sémantiques que les personnes possèdent sur les informations qu'elles doivent traiter et retenir peuvent parfois influencer leur capacité à auto-évaluer l'exactitude de leur mémoire. Comme anticipé, la manipulation de la congruence sémantique des cibles et des distracteurs dans le test de reconnaissance n'a pas eu d'effet sur la performance de premier ordre. Cela peut être expliqué par le fait que la congruence a été manipulée lors du test et n'a donc pas pu jouer sur la force de la trace mnésique. De plus, la congruence des images n'était pas pertinente pour la réponse de type-1 car la question portait sur les items ayant été vus dans la phase d'encodage. Ainsi, il n'y avait pas de raison de penser que la congruence allait jouer sur les performances mnésiques.

En revanche, la congruence des images a eu un faible effet sur la magnitude des RCJ. Contrairement à ce que nous attendions, les participants n'ont pas été influencés par la différence de congruence entre la cible et le distracteur pour faire les RCJ. Les items présentés dans des contextes incongruents sont généralement identifiés moins rapidement que dans un contexte congruent (LaPointe et al., 2013; Ortiz-Tudela et al., 2017), ce qui laisse supposer que les items sont moins bien traités dans un contexte incongruent et donc que la fluence de traitement d'un item est plus faible par rapport à un contexte congruent. Les participants ne semblent pas avoir utilisé la différence relative entre la fluence de traitement de la cible et celle de son distracteur comme indice pour faire les RCJ. La magnitude des RCJ a plutôt été influencée par la congruence des distracteurs puisque les participants avaient plus confiance en l'exactitude de leur réponse lorsque le distracteur était congruent que lorsqu'il était incongruent.

Afin de voir si ce résultat pouvait être expliqué par des différences dans le traitement des deux stimuli (i.e., cible et distracteur) en général (plutôt que par comparaison), nous avons analysé les TR dans chacune des conditions. Globalement, la magnitude des RCJ était plus élevée pour les réponses données rapidement que pour celles données plus lentement. Ce résultat est en accord avec la littérature qui a montré que les jugements rétrospectifs semblent majoritairement influencés par la fluence de réponse (i.e., *answer fluency*; Costermans et al., 1992; Kelley & Lindsay, 1993). Ce résultat n'est pas spécifique aux tâches de mémoire et a également été observé sur des tâches de résolution de problèmes (e.g., Ackerman & Zalmanov, 2012) ou de décisions perceptives (e.g., Kiani et al., 2014). Les TR étaient légèrement influencés par la congruence des distracteurs. Les participants ont été plus rapides pour répondre dans les essais avec des distracteurs congruents que dans les essais avec des distracteurs incongruents. De plus, l'effet de la congruence des distracteurs sur la magnitude des RCJ disparaissait lorsque les TR étaient pris en compte. Ainsi, les temps de réponses semblent expliquer l'effet de la congruence du distracteur sur la magnitude des jugements. En effet, les participants semblent avoir donné des RCJ plus élevés dans les essais avec des distracteurs congruents car ils étaient légèrement plus rapides pour répondre. Ce résultat corrobore l'idée que les RCJ sont influencés par la fluence de réponse (Kelley & Lindsay, 1993), un indice qui semble être domaine-général.

Une explication possible pour expliquer ces différences de TR est que lorsque les participants se souvenaient de la cible vue (i.e., essais corrects), sa congruence n'avait peu d'importance car les participants étaient de toute manière plus familiers avec la cible qu'avec son distracteur. Sur une tâche de reconnaissance à deux alternatives, une bonne réponse peut être sélectionnée grâce la familiarité avec la cible ou sa recollection (i.e., acceptation de la cible), mais également parce que la personne est certaine de ne pas avoir vu le distracteur (i.e., rejet du distracteur, *recall-to-reject*; Rotello et al., 2000). Ainsi, dans les essais corrects, on peut faire l'hypothèse que les distracteurs congruents étaient mieux traités (voir identifiés) que les distracteurs incongruents, ce qui aurait pu les aider à faire un choix sur le fait de ne pas avoir vu l'item et donc à rejeter cette réponse plus rapidement.

Comparer les RCJ aux performances de premier-ordre en analysant le biais métacognitif, a permis de mettre en évidence que pouvoir bien traiter à la fois la cible et le distracteur a amené les participants à surestimer légèrement leur performance mnésique. La magnitude des RCJ ayant été influencée par la congruence sémantique sans que la performance de premier ordre le soit, cette expérience fournit un exemple d'illusion métacognitive créée par la manipulation de facteurs sémantiques sur les RCJ. Ainsi, les connaissances préalables que les individus ont sur le monde (e.g., savoir qu'un objet est généralement rencontré dans tel ou tel contexte), influence la manière dont les personnes

évaluent leurs souvenirs et peuvent donc parfois les amener à se tromper. Ces résultats semblent être expliqués par le fait que leurs réponses de type-1 étaient plus rapides dans le cas où les distracteurs étaient congruents, bien que leurs réponses n'aient pas pour autant été plus correctes.

Les résultats n'ont pas mis en évidence l'influence de la congruence sémantique sur la sensibilité métacognitive, contrairement à ce que nous attendions. L'effet de la congruence du distracteur était significatif lorsque l'AUROC2 était utilisée comme mesure mais l'analyse bayésienne complémentaire n'a pas permis de confirmer ce résultat. De plus, les analyses sur l'efficacité métacognitive, qui est une mesure de sensibilité pour laquelle la performance de type-1 est contrôlée, n'ont pas révélé d'effet de la congruence sémantique. Il est donc possible que les résultats obtenus avec les AUROC2 à l'aide des statistiques fréquentistes soient un faux positif. Ainsi, la manipulation de la congruence sémantique ne semble pas avoir affecté la capacité des personnes à discriminer les essais réussis des essais échoués.

### 7.3. **Expérience 7.2 : Effet de congruence sémantique sur les prédictions de reconnaissance en mémoire épisodique**

Dans cette étude, la congruence sémantique a été manipulée à l'encodage. Le but était de voir comment les participants ajustaient leurs PoR lorsque la congruence sémantique pouvait influencer la performance de premier ordre. Pour cela, les participants ont dû mémoriser des images d'items présentés dans des scènes congruentes ou incongruentes et réaliser un paradigme de PoR. Nous faisons l'hypothèse que les participants retiendraient mieux les items congruents qu'incongruents (Ortiz-Tudela et al., 2017). Si l'indice le plus utilisé était la familiarité avec la scène vue au moment de la PoR, nous ne prédisions pas de différence de magnitude entre les essais congruents et incongruents (voir Ortiz-Tudela et al., 2017 pour une absence de différence de mémoire pour les cibles congruentes et incongruentes). Cependant, si les participants utilisaient également des processus mnésiques plus épisodiques (i.e., récupération, même partielle, de la cible), nous prédisions une plus grande magnitude dans les essais avec une cible congruente que pour ceux avec une cible incongruente. Aucune étude n'ayant évalué l'effet de la fluence conceptuelle sur des PoR, faites au moment de la récupération en mémoire, l'estimation de la sensibilité métacognitive dans cette expérience était exploratoire.

#### 7.3.1. **Méthode**

##### **Participants**

Cent-huit participants de nationalité américaine et de langue maternelle anglaise<sup>36</sup> ont été recrutés sur Prolific pour cette étude ( $M_{age} = 30.89$ ,  $ET_{age} = 6.33$ , 56 femmes). Ce nombre a été déterminé sur la base d'une analyse de puissance réalisée sur G\*Power pour observer un effet faible à moyen de 0.30, une puissance de .80 et un seuil  $\alpha = .05$ , et en prenant en compte qu'environ 20% de l'échantillon a été exclu des analyses dans l'Expérience 7.1. Nous avons exclu 22 participants des analyses (voir section [Résultats](#) pour le détail). Ainsi, les données de 86 participants ont été analysées ( $M_{age} = 30.64$ ,  $ET_{age} = 5.65$ , 47 hommes).

##### **Protocole**

L'expérience a été réalisée sur la plateforme Gorilla et était décomposée en 3 phases (voir Figure 7.5). Dans la phase d'encodage, une image de scène visuelle était présentée seule pendant 750ms puis un item (i.e., animal, objet ou aliment) apparaissait dedans pour 250ms supplémentaires. Les participants devaient apprendre 138 images scène-item. La moitié de

---

<sup>36</sup> La majorité des participants français inscrits sur la plateforme Prolific avait participé à l'étude pilote de l'expérience 7.1 ou à l'expérience 7.1. Un bien plus grand nombre d'américains (USA) étant inscrit sur la plateforme, nous avons changé les critères de sélection et recruté des participants américains pour faciliter le recrutement.

ces items était sémantiquement congruente (e.g., voilier dans la mer ; i.e., conceptuellement fluent) et l'autre moitié était incongruente (e.g., baleine dans un champs ; i.e., conceptuellement non fluent) avec la scène visuelle. Dans la seconde phase, les participants réalisaient une tâche distractive de détection de cibles similaire au paradigme de Posner (1980), identique à celle de l'Expérience 7.1. Pour essayer de limiter encore plus les effets plafonds sur la tâche de reconnaissance, les participants devaient également faire des calculs mentaux entre certains essais de la tâche distractive. Enfin, dans une troisième phase, 92 des images de scènes seules vues dans la phase d'apprentissage étaient présentées une par une et pour chacune, les participants devaient prédire sur une échelle de 50% à 100% à quel point ils étaient confiants qu'ils pourraient reconnaître l'item qui était apparu dedans (PoR). Après chaque prédiction, ils devaient sélectionner le bon item parmi deux propositions de réponses. Nous avons augmenté le nombre d'images à l'encodage (138) par rapport au nombre d'essais au test (92) afin d'augmenter la difficulté de la tâche d'encodage et éviter des performances plafonds sur la tâche de premier ordre.

Pour éviter d'autres effets confondus tels que l'effet de fréquence des items, les cibles et les distracteurs étaient contrebalancés à travers les participants : les cibles pour l'un étaient les distracteurs pour d'autres, et inversement. À noter que pour certaines paires cible-distracteur, la cible n'était pas assez (in)congruente dans l'image correspondante du distracteur. Ainsi, afin de s'assurer que la différence de congruence de la cible et du distracteur dans la scène soit réduite, le distracteur 1 d'une cible 1 pour un participant A n'avait pas toujours cette même cible 1 comme distracteur pour un participant B mais un autre item 2 qui servait de cible 2 pour A. Dans ce cas, la cible 1 pour le participant A était utilisée comme distracteur d'une cible 3 pour le participant B. En résumé, chaque item était vu dans une scène congruente pour certains, dans une scène incongruente pour d'autres, ou comme distracteurs.

### **Matériel**

Deux-cents images ont été sélectionnées pour cette étude. La plupart étaient issues des images utilisées dans l'[Expérience 7.1](#). Cependant, pour éviter que la congruence soit utilisée comme information pour répondre à la tâche de reconnaissance, les distracteurs ont été choisis de manière à ce qu'ils soient de même congruence avec la scène que les cibles (e.g., incongruent si cible incongruente). En effet, nous ne voulions pas que les participants répondent à la tâche juste en se rappelant que l'item était congruent plutôt qu'en se rappelant de l'item en soit. Les 100 paires cibles-distracteurs ont été créées afin que les deux soient congruentes dans la scène congruente de l'une et de l'autre, et incongruente dans la scène incongruente de l'une et l'autre. Ainsi, certaines scènes de fond ont été changées par rapport à l'expérience 7.1 afin d'augmenter la (in)congruence du distracteur dans la scène de la cible pour se rapprocher du degré de (in)congruence de la cible dans cette scène.



**Encart 3. Description des études pilotes pour la sélection des stimuli de l'expérience 7.2**

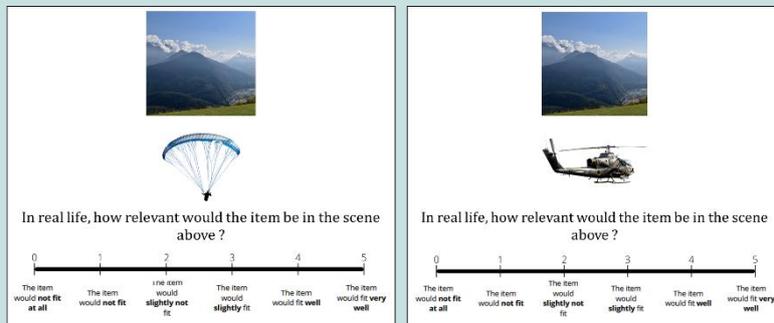
**Pilote 1. Vérification de la congruence chez des américains**

Dans une première étude, 73 participants américains âgés de 19 à 39 ans ( $M_{age} = 28.68$ ,  $ET_{age} = 6.41$ ; 33 femmes) recrutés en ligne via Prolific ont vu des montages scènes-items. Ils devaient juger sur une échelle de 0 (l'item ne va pas du tout) à 5 (l'item va très bien dans la scène) à quel point l'item était congruent avec la scène. Ce premier pilote permettait de vérifier que les items étaient bien perçus comme congruents ou incongruents dans la phase d'encodage. Le protocole était similaire à celui utilisé pour l'étude pilote de l'Expérience 7.1 réalisée auprès de français. Cependant, puisque nous projetions de réaliser l'Expérience 7.2 auprès d'américains, nous souhaitions vérifier que les items étaient toujours perçus comme (in)congruent en cas de différence culturelle. Chaque participant voyait un tiers des montages. Nous avons repris quelques items ayant été exclus pour l'expérience 7.1 et changé leur contexte afin qu'ils soient plus (in)congruents.

Les mêmes analyses que pour le pilote de l'expérience 7.1 ont été conduites (voir **Encart 2.** ). Puisque certaines images n'étaient pas perçues comme étant assez (in)congruentes qu'espéré, nous les avons exclues de l'étude. En effet, quatre images supposée congruentes ont été jugées en moyenne sous 2.6/5 (paon, aspirateur, canne, et machine à écrire). Les images congruentes étaient bien en moyenne jugées comme étant plus congruentes ( $M = 4.15$ ,  $ET = 0.62$ ) que les images incongruentes ( $M = 0.30$ ,  $ET = 0.31$ ) ( $t(357) = 104.76$ ,  $p < .001$ ).

**Pilote 2. Estimation de la congruence des cibles et des distracteurs**

Dans une seconde étude, 155 participants américains âgés entre 18 et 40 ans ( $M_{age} = 30.46$ ,  $ET_{age} = 6.09$ ; 68 femmes) ont été recrutés en ligne via Prolific. Chacune des 400 scènes (200 congruentes et 200 incongruentes) étaient présentées deux fois : une fois avec la cible associée (voir image gauche ci-dessous) et une fois avec son distracteur (voir image droite ci-dessous). Les images de la scène et de l'item (i.e., cible ou distracteur) étaient présentées séparément sur le même écran mais l'item n'était pas inséré à l'intérieur de l'image. Toutes les scènes n'étaient pas vues par les mêmes participants. Chaque participant ne voyait qu'une scène des paires cible-distracteur (la congruente ou l'incongruente de l'un des deux items de la paire) pour éviter un jugement comparatif (i.e., jugement en fonction d'une autre vue avant) et avoir un jugement plus absolu. Les participants jugeaient si l'item irait bien dans la scène dans la vie réelle sur une de 0 (l'item n'irait pas du tout) à 5 (l'item irait très bien).



Sur la base de ces données, nous avons exclu 4 paires supplémentaires car la différence des estimations moyennes pour la cible et le distracteur était en moyenne supérieure à 2 points sur l'échelle pour au moins une des scènes. En effet, pour ces paires et pour la scène congruente d'un des items, les distracteurs avaient reçu des jugements trop faibles comparés aux jugements reçus par les cibles. Ainsi, 92 paires (368 images : une image congruente et une image incongruente pour chaque item des 92 paires) ont été gardées pour l'Expérience 7.2. Les cibles étaient moyennes perçues comme plus congruentes dans leur scène respective ( $M = 4.29$ ,  $ET = 0.57$ ) que les distracteurs ( $M = 4.17$ ,  $ET = 0.63$ ) ( $t(184) = 2.33$ ,  $p = .02$ ). De même, les cibles étaient en moyenne perçues comme plus incongruentes dans leur scène respective ( $M = 0.24$ ,  $ET = 0.23$ ) que leurs distracteurs ( $M = 0.33$ ,  $ET = 0.33$ ) ( $t(185) = 3.39$ ,  $p < .001$ ). Cependant, pour certaines paires, les distracteurs étaient en moyenne perçues comme plus (in)congruent que les cibles. Ainsi, dans le test de reconnaissance de l'expérience 7.2 finale, les participants ne pouvaient pas utiliser la congruence des deux items avec la scène présentée pour deviner celui qu'ils avaient appris.

### 7.3.2. Résultats

#### **Exclusion des données**

Nous avons exclu les données de 9 participants car ils ont eu des performances globales à la tâche de reconnaissance au niveau du hasard (<55%). Nous avons également exclu 6 participants en raison de performances plafonds (> 95 %CR) sur la condition congruente et 4 participants supplémentaires qui n'avaient pas assez varié leurs prédictions puisque ces deux cas de figure empêchent le calcul de score de sensibilité métacognitive. Enfin, les données de 3 participants supplémentaires ont été exclues car ils présentaient des scores extrêmes sur au moins une des mesures dans une condition (1 sur la magnitude, 2 sur les scores gamma). Pour plus de cohérence, nous avons exclu ces données de toutes les analyses.

#### **Performance de mémoire (%RC)**

L'effet de la congruence des scènes sur la reconnaissance des items était significatif ( $t(85) = 4.05, p < .001, d_z = 0.44$ ). Les participants ont mieux reconnu les items appris dans un contexte congruent ( $M = 0.76, SD = 0.10$ ) que dans un contexte incongruent ( $M = 0.73, SD = 0.10$ , voir Figure 7.6a).

#### **Magnitude et biais métacognitif**

Les prédictions étaient également significativement influencées par la congruence des scènes ( $t(85) = 6.24, p < .001, d_z = 0.67$ )<sup>37</sup>. Les participants prédisaient de meilleures performances pour les items vus dans des scènes congruentes ( $M = 0.67, SD = 0.09$ ) que pour ceux appris dans des scènes incongruentes ( $M = 0.65, SD = 0.08$ , voir Figure 7.6b pour proportions). Les analyses n'ont pas révélé d'effet significatif de la congruence sur le biais métacognitif ( $t(85) = 1.81, p = .07, d_z = 0.20$ , voir Figure 7.6c)<sup>38</sup>. Une analyse bayésienne complémentaire (t-test pour mesures répétées, non pré-enregistrée) a révélé une preuve anecdotique pour l'hypothèse nulle lorsque l'effet de congruence sur le biais était testé ( $BF_{01} = 1.76$ ).

#### **Sensibilité métacognitive**

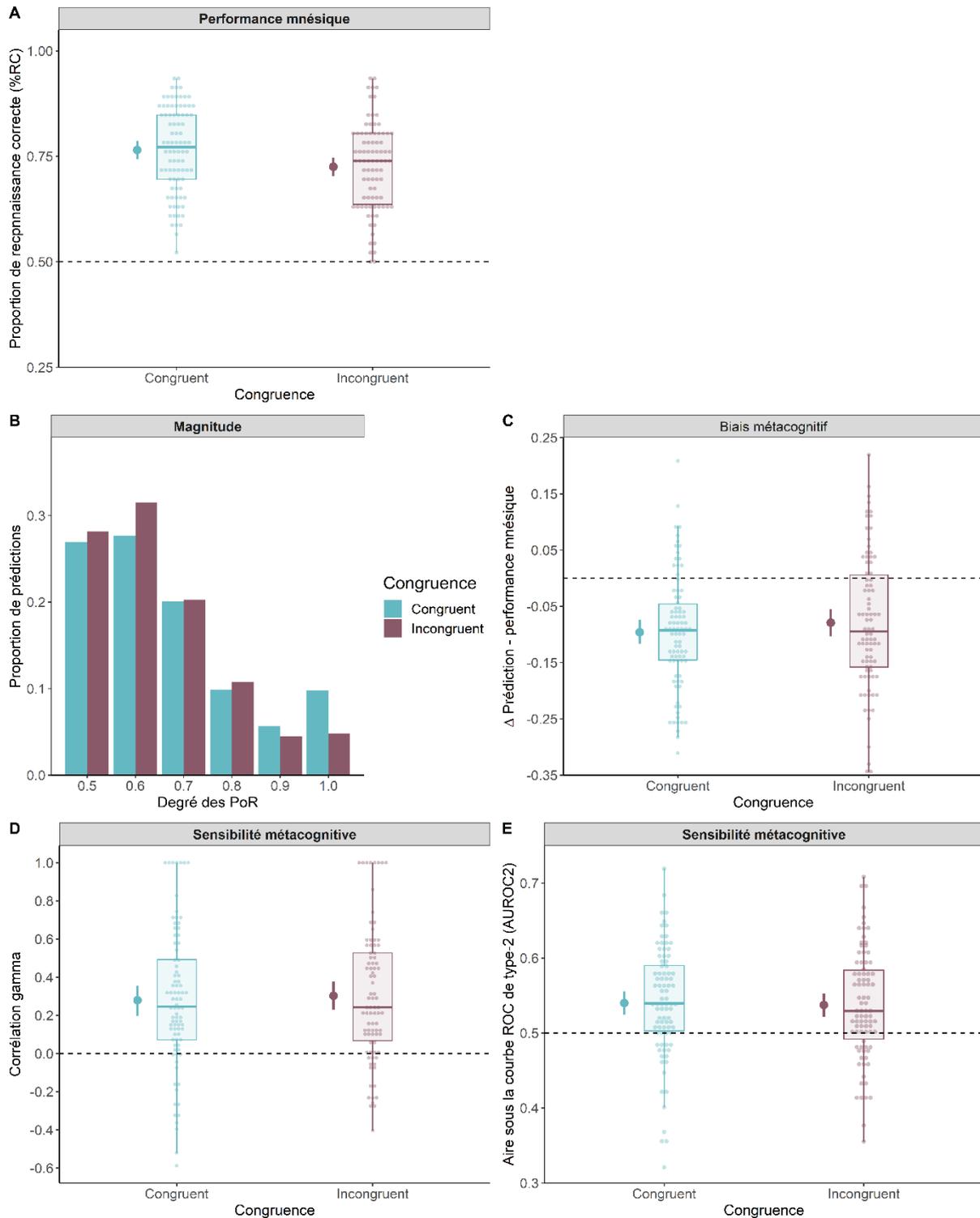
La sensibilité métacognitive n'a pas été significativement influencée par la congruence des scènes dans lesquelles les items étaient présentés, que ça soit en l'évaluant avec les scores gamma ( $t(85) = 0.47, p = .64, d_z = 0.05$ , voir Figure 7.6d) ou les aires sous la courbe ROC de type-2 ( $t(85) = 0.25, p = .80, d_z = 0.03$ , voir Figure 7.6e). Des analyses bayésiennes

---

<sup>37</sup> Les résidus du modèle n'étaient pas distribués normalement. En utilisant un test paramétrique de Wilcoxon :  $V = 2497.5, p < .001$ .

<sup>38</sup> Les résidus du modèle n'étaient pas distribués normalement. En utilisant un test paramétrique de Wilcoxon :  $V = 1549, p = .17$ .

complémentaires (non pré-enregistrées) ont fourni des preuves modérées de ces effets nuls (sur les scores gammas:  $BF_{01} = 7.56$  ; sur les AUROC2:  $BF_{01} = 8.15$ ).



**Figure 7.6 :** Résultats de l'expérience 7.2.

A) Performance de premier ordre, B) Magnitude des RCJ, C) Biais métacognitif, D) Scores gamma, E) AUROC2. Les points épais représentent les moyennes par condition, avec l'intervalle de confiance à 95%. Les points plus petits sont les données pour chaque participant.

### **7.3.3. Discussion de l'expérience 7.2.**

L'objectif de cette expérience était de manipuler la congruence sémantique d'images à l'encodage afin de voir si cela avait un impact sur la mémoire, mais surtout si les participants arrivaient à ajuster leurs PoR en conséquence. Les résultats ont montré que les participants ont plus reconnu les items congruents que les items incongruents. Cet effet de congruence sémantique sur la mémoire épisodique, a auparavant été observé avec différents types de stimuli (e.g., visages, Atienza et al., 2011 ; mots, Packard et al., 2017, 2020; Schulman, 1974). Certains proposent qu'il pourrait découler d'un processus d'assimilation par l'activation des connaissances antérieures et de réseaux neuronaux préexistants, qui permettrait de booster l'encodage en mémoire (van Kesteren et al., 2012).

De plus, lorsque les participants les scènes seules, au moment de la prédiction, ils ont utilisé la congruence de l'item qui avait été présenté dedans à l'encodage comme indice puisque la magnitude des PoR était plus élevée dans les essais congruents que dans les essais incongruents. Ainsi, contrairement à Mazancieux (2020, Chapitre 6, expérience 3), notre étude montre que la manipulation de la fluence conceptuelle peut avoir un impact sur les PoR. Pour rappel, dans leur étude (Mazancieux, 2020), les participants devaient mémoriser des patterns visuels abstraits insérés dans des mots (i.e., essais fluents) ou des pseudo-mots (i.e., essais non-fluents). Même si la fluence conceptuelle n'était pas pertinente pour la tâche de premier ordre, la performance de premier ordre était meilleure dans les essais fluents que non-fluents. En revanche, la magnitude des PoR n'était pas affectée par cette manipulation.

Les résultats de notre étude montrent également que le biais métacognitif et la sensibilité ne sont pas impactés par la congruence sémantique. Ainsi, cela suggère que les participants ont bien pu prédire leur performance et donc que la congruence sémantique semble être un indice diagnostique utilisé par les participants. Cette seconde expérience montre que les connaissances que les personnes ont sur les stimuli à traiter à l'encodage a des conséquences à long terme (i.e., pas seulement sur des aspects perceptifs) : ces connaissances influencent la mémoire de ces stimuli et les participants semblent être capables d'ajuster leurs jugements métacognitifs en conséquence.

Des travaux supplémentaires semblent être nécessaires afin de déterminer l'origine de l'effet de congruence sur la magnitude des PoR. En effet, il est possible que les participants aient été plus capables de rappeler les cibles au moment de faire leurs PoR dans les essais congruents par rapport aux essais incongruents, et que des informations sur les cibles congruentes aient été plus accessibles. Les travaux menés sur le FOK, qui sont un exemple particulier de PoR, montrent que les FOK sont influencés par la familiarité avec l'indice et la récupération d'information sur la cible (Koriat & Levy-Sadot, 2001; Thomas et al., 2012). De plus, de précédentes études (Ortiz-Tudela et al., 2017) n'ont pas permis de mettre en avant

que la mémoire des scènes est accrue dans les cas où un item congruent est présenté par rapport à des cas avec des items incongruents. Il est donc plus probable que les facteurs liés aux cibles (i.e., items) et non aux indices (i.e., scènes) aient contribué à l'augmentation de la magnitude des PoR dans les essais congruents par rapport aux essais incongruents. Ainsi, d'autres études permettraient de voir si, dans les essais congruents, les cibles étaient plus accessibles (i.e., déjà récupérées en mémoire au moment de la prédiction donc très prédictif d'une reconnaissance future), si des informations partielles étaient plus accessibles (e.g., les participants ne se rappelaient plus de l'item mais se souvenaient qu'il était congruent), ou si d'autres facteurs jouaient.

Enfin, les performances mnésiques ayant été affectées par la manipulation de la congruence sémantique, il pourrait également être intéressant de créer une condition dans laquelle les essais incongruents seraient mieux retenus que les essais congruents (e.g., par l'augmentation du temps de traitement à l'encodage ou par un critère de rappel ; voir Carroll et al., 1997). Il serait alors possible de voir si les informations sémantiques sur les stimuli à traiter sont utilisés comme indice et influence la magnitude des PoR même si la performance de premier ordre est affectée différemment, ou si les PoR sont majoritairement influencés par la force de la trace mnésique.

#### **7.4. Discussion générale du Chapitre 7**

Les deux études montrent que la congruence sémantique d'images à traiter peut influencer les jugements métacognitifs faits sur une tâche de mémoire épisodique, que ces jugements soient rétrospectifs (Expérience 7.1) ou prédictifs (Expérience 7.2). Ces résultats peuvent être mis en perspective des précédentes études portées sur la fluence conceptuelle dans le JOL (e.g., Castel et al., 2007; Connor et al., 1997; Dunlosky & Matvey, 2001; Undorf & Erdfelder, 2015) et sur les RCJ (Mazancieux, 2020).

L'expérience 7.1 a montré qu'il était possible de créer une illusion métacognitive en manipulant des facteurs sémantiques telles que la congruence sémantique d'images. Cette expérience a permis de mettre en évidence que les connaissances sémantiques peuvent amener les participants à modifier leurs jugements métacognitifs même lorsque leur performance mnésique ne change pas. L'expérience 7.2 a mis en évidence que les PoR peuvent être influencés par la fluence conceptuelle lorsqu'elle est manipulée par la congruence sémantique d'images, notamment dans le cas où les performances en mémoire sont également affectées par cette manipulation. Ainsi, les participants sont capables d'ajuster leurs prédictions en accord avec leurs performances mnésiques futures. En revanche, la sensibilité métacognitive n'était pas influencée par la manipulation de la congruence

sémantique, que ça soit pour les jugements rétrospectifs (Expérience 7.1) ou pour les jugements prédictifs (Expérience 7.2).

D'autres études sont nécessaires afin de déterminer l'origine de l'effet de congruence sémantique sur la magnitude des jugements métacognitifs. En effet, les processus sous-tendant l'effet de la fluence de traitement sur les jugements métacognitifs sont encore débattus (Mueller et al., 2013; Unkelbach & Greifeneder, 2013). Certains suggèrent que des processus basés sur l'expérience subjective entraînent ces effets de fluence (Undorf et al., 2017; Undorf & Erdfelder, 2015). Selon ces travaux, les processus de fluence seraient directs : la sensation subjective de facilité de traitement amènerait à un jugement métacognitif élevé par un mécanisme d'attribution (e.g., c'est facile à traiter donc je suis plus confiant que c'est la bonne réponse). D'un autre côté, d'autres travaux indiquent que les effets de fluence conceptuelle, comme l'effet de *relatedness* sur les JOL, découlent de croyances métacognitives, et donc de processus basés sur la théorie (i.e., *theory-based*). Notamment, les individus auraient des *a priori* sur l'effet qu'aurait une manipulation sur leurs performances et modifieraient donc leurs jugements métacognitifs en conséquence (Mueller et al., 2013). Dans l'exemple de l'effet de *relatedness*, les individus auraient une croyance explicite sur la manière dont le degré d'association entre deux mots affecte la mémoire. Il est également possible que les deux types de processus (i.e., expériences subjectives et croyances) jouent un rôle complémentaire dans les mêmes jugements (Oppenheimer, 2008). Ainsi, dans les deux expériences présentées ici, il est possible que des processus basés sur l'expérience aient entraîné un ajustement des jugements en fonction de la congruence sémantique. Cette influence pourrait être indirecte, dû à une médiation par la fluence de réponse dans l'expérience 7.1, ou plus en lien avec des facteurs mnésiques dans l'expérience 7.2. Pour les RCJ, ce serait donc un indice plus domaine-général, la fluence de réponse (voir Ackerman & Zalmanov, 2012; Kelley & Lindsay, 1993; Kiani et al., 2014 pour des exemples dans différents domaines), qui interviendrait : fluence de traitement au moment du test qui amène les participants à répondre rapidement et à avoir plus confiance en leur réponse. Pour les PoR, la fluence de traitement jouerait sur l'encodage des informations et sur des processus mnésiques de rétention ou de récupération, et les indices mnésiques seraient ensuite utilisés pour faire les PoR. Comme cela déjà été montré (Undorf et al., 2018; Undorf & Bröder, 2020), plusieurs indices peuvent être utilisés pour un même jugement. Nous ne pouvons pas non plus exclure que des processus basés sur la théorie (i.e., croyances naïves) aient également joué un jeu dans les effets de congruence sémantique observés.

## CHAPITRE 7 : CONCLUSION

Dans deux expériences, nous avons montré que les informations stockées en mémoire à long terme même sémantiques peuvent servir d'indices aux jugements métacognitifs faits sur une tâche épisodique, et parfois amener les participants à se tromper dans leur évaluation.

La **fluence conceptuelle** influence les RCJ et les PoR faits sur un test de mémoire épisodique. En effet, lorsque les informations à traiter dans le test de mémoire ne sont pas congruentes avec les connaissances sémantiques, les personnes avaient moins confiance en l'exactitude de leurs réponses, sans que cela n'affecte leur sensibilité métacognitive et bien que leur performance mnésique ne soit pas impactée. Cet effet semble venir d'une réduction de la **fluence de réponse** dans les essais où les distracteurs étaient incongruents, sûrement dû à une plus grande difficulté à traiter ces items. De même, les participants prédisaient reconnaître moins d'items ayant été appris dans un contexte incongruent, ce qui reflétait le pattern de performance mnésique.

Ainsi, la **mémoire sémantique** est utilisée pour le monitoring en mémoire épisodique, même lorsque cela n'est pas toujours pertinent.

# Chapitre 8.

## Discussion générale et perspectives

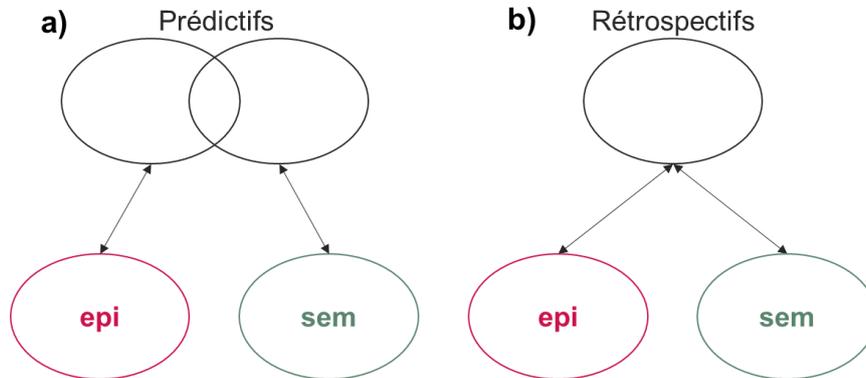
---

En résumé, la métacognition regroupe un ensemble de compétences, dont le monitoring qui représente entre autres l'aptitude à pouvoir juger a posteriori de l'exactitude de ses souvenirs ou encore de pouvoir prédire sa capacité future à récupérer des informations en mémoire ([Chapitre 1](#)). Selon la théorie de l'utilisation d'indices (Koriat, 1997), les personnes infèrent l'exactitude d'un souvenir ou d'une connaissance sur la base d'indices de différentes natures (intrinsèques, extrinsèques, ou mnésiques ; basés sur des expériences subjectives ou des théories, [Chapitre 2](#)). Un ensemble de travaux en psychologie expérimentale, en neuro-imagerie et en neuropsychologie ont eu pour but de comprendre si les indices utilisés et les mécanismes mis en jeu pour les jugements de monitoring sont domaine-généraux ou domaine-spécifiques. L'idée de domaine-généralité suppose que les mêmes mécanismes sont mis en jeu dans les jugements métacognitifs peu importe le domaine cognitif évalué dans la tâche de premier ordre (i.e., mémoire, perception). À l'inverse, une domaine-spécificité suggère des mécanismes propres à chacun des domaines cognitifs évalués.

Ces mêmes questions se posent au sein du champ de la métamémoire vis-à-vis de la forme de mémoire évaluée par la tâche de premier ordre (i.e., épisodique ou sémantique), et selon le moment où le jugement métacognitif est fait (avant ou après la réponse de type-1). De précédents travaux (voir [Chapitre 2](#)) suggèrent que les jugements rétrospectifs utilisent des indices et processus domaine-généraux. D'un autre côté, les jugements prédictifs faits avant la réponse de type-1 comme le FOK seraient plus mémoire-spécifiques, c'est-à-dire qu'ils se baseraient sur des indices spécifiques à la forme de mémoire testée dans la tâche de premier ordre. Notons cependant que nous considérons quelques points communs entre les jugements prédictifs épisodiques et sémantiques, comme l'appui sur les fonctions exécutives, et ne proposons donc pas une distinction stricte entre métacognition prédictive épisodique et métacognition prédictive sémantique. Nous proposons d'intégrer les résultats obtenus dans cette thèse dans un modèle théorique inspiré de la littérature traitant de la généralité/spécificité de la métacognition (voir Figure 8.1).

Pour rappel, nous avons deux objectifs principaux pour mieux comprendre le rôle de la mémoire dans le monitoring. Tout d'abord, nous avons cherché à approfondir les connaissances sur la distinction entre métacognition épisodique et sémantique en étudiant les performances métacognitives de personnes présentant un déclin de la mémoire épisodique

(Chapitre 4) ou sémantique (Chapitre 5). Dans un second temps, nous avons pris en compte l'interaction entre les processus de premier ordre épisodiques et sémantiques dans le fonctionnement normal et exploré l'intervention des connaissances sémantiques dans la métacognition épisodique (Chapitres 6 et 7).



**Figure 8.1** : Proposition d'un modèle théorique récapitulant les résultats de la littérature contribuant à la compréhension des relations entre processus de premier ordre (en bas) et de second ordre (en haut) selon que les jugements métacognitifs soient prédictifs (à gauche) ou rétrospectifs (à droite) (schéma inspiré de Mazancieux et al., soumis).

a) Les jugements prédictifs seraient plus mémoire-spécifiques. Les indices et mécanismes sous-jacents dépendraient en partie de la forme de mémoire testée sur la tâche de premier ordre. Ici, les jugements prédictifs regroupent les prédictions de rappel ou de reconnaissance pouvant être faits à la fois sur les tâches épisodiques et les tâches sémantiques (e.g., PoR, FOK).

b) Les jugements rétrospectifs seraient plus domaine-généraux.

### 8.1. Résumé des résultats

Plusieurs aspects du rôle de la mémoire dans les jugements métacognitifs ont été abordés dans ce travail, et plusieurs méthodes ont été utilisées afin de vérifier nos hypothèses et de répondre à notre problématique. Les principaux résultats sont présentés dans le Tableau 8.1.

Tout d'abord, les deux chapitres présentant des études sur la considération d'une dissociation entre métacognition épisodique et sémantique ont amené à des conclusions différentes mais pas pour autant contradictoires (Chapitre 4 et 5). En réalisant une méta-analyse, dans le Chapitre 4, nous avons pu confirmer une dissociation de la sensibilité métacognitive dans le FOK chez les personnes âgées. En effet, les travaux publiés montraient une réduction de la sensibilité du FOK avec l'âge lorsque la tâche de premier ordre testait la mémoire épisodique mais non lorsque cette tâche portait sur la mémoire sémantique. Ce résultat était congruent avec les études ayant évalué le eFOK et le sFOK chez les mêmes personnes (Douchemane et al., 2007; Morson et al., 2015; Souchay et al., 2007). Le pattern de dissociation épisodique/sémantique de second ordre reflétait le pattern de performance de premier ordre. En analysant plus particulièrement des données issues de notre laboratoire, nous avons pu évaluer plus précisément le rôle des performances mnésiques dans le déclin

du eFOK. Les différences de sensibilité liées à l'âge étaient réduites lorsque les performances de rappel et de reconnaissance étaient contrôlées ou lorsque ces performances étaient rendues égales entre les deux groupes d'âge. L'ensemble des résultats montrait que le déclin spécifique du eFOK dans le vieillissement découlait d'une réduction des capacités mnésiques épisodiques.

Ainsi, le Chapitre 4 apporte des preuves additionnelles à celles qui ont déjà montré qu'un déclin de la mémoire épisodique peut entraîner une diminution de la sensibilité des prédictions épisodiques, comme c'est également le cas dans la MA (Ernst et al., 2015). À notre connaissance, aucune étude n'avait auparavant évalué avec des mesures de métacognition locale, les capacités métacognitives de personnes présentant un trouble mnésique plus sémantique. Dans le Chapitre 5, les capacités métacognitives de personnes atteintes de DS (i.e., une maladie neurodégénérative entraînant une dégradation du système sémantique) ont été évaluées sur des tâches épisodiques et sémantiques, par le biais de jugements prédictifs et rétrospectifs. Les résultats de ce chapitre n'ont pas permis de mettre en évidence un déficit métacognitif général dans cette pathologie, ni spécifique aux tâches sémantiques. Cela était le cas concernant leur capacité à prédire ou à évaluer rétrospectivement leurs souvenirs et connaissances. Aucune double dissociation n'a été observée en comparant les performances de ces patients à ceux présentant une MA. Ainsi, nos résultats ne montrent pas qu'un déficit de premier ordre sémantique est lié à un déficit de sensibilité sémantique. Ils ne suggèrent donc pas un rôle primordial de la préservation du système sémantique pour les jugements métacognitifs. Cependant, la réduction tendancielle de sensibilité du eRCJ chez les DS nécessite d'être investiguée davantage.

Dans les deux derniers chapitres expérimentaux, nous avons exploré l'utilisation des informations sémantiques pour les jugements métacognitifs en mémoire épisodique dans le fonctionnement normal. Dans le Chapitre 6, dans 4 études, les participants ont appris des paires d'items qui avaient une signification (i.e., mots réels) ou qui n'avaient pas de représentation stockée en mémoire sémantique (i.e., pseudo-mots). Les personnes utilisaient à la fois les informations propres aux indices des paires et celles liées aux cibles à récupérer comme indice pour prédire une récupération en mémoire future, même si ces informations n'affectaient pas toujours la performance de premier ordre tel que prédit. La sensibilité métacognitive était plus faible lorsqu'au moins un des items à traiter ne possédait pas de représentation sémantique. Cela montre que le traitement sémantique est important pour avoir une sensibilité métacognitive optimale. Cet effet ne semblait pas venir de facteurs liés à l'encodage des items. Cependant, ce résultat disparaissait lorsque seuls les items non-rappelés étaient considérés. Ainsi, nous avons discuté de la potentielle médiation de l'influence des traitements sémantiques sur la sensibilité par l'influence de la mémoire sémantique sur les capacités de récupération en mémoire épisodique.

Plusieurs travaux ont montré que la fluence conceptuelle, manipulée sur la base des connaissances générales, pouvait être utilisée comme indice pour juger ses performances en mémoire épisodique avec du matériel verbal (Mazancieux, 2020; Mueller et al., 2013). Or, le système sémantique intervient également dans le traitement d'images. Ainsi, le Chapitre 7, visait à étudier l'effet de la congruence sémantique d'images sur les jugements métacognitifs prédictifs et rétrospectifs. Deux études ont montré que les connaissances sémantiques étaient utilisées pour les jugements rétrospectifs et prédictifs. En effet, les participants étaient moins confiants en l'exactitude de leurs réponses dans les cas où les distracteurs étaient présentés dans un contexte incongruent au moment du test de mémoire, sans que leur performance mnésique ne soient pour autant impactée, créant ainsi une illusion métacognitive. Cet effet semblait découler d'un ralentissement dans les réponses dans cette condition. Cela suggérait une médiation de l'effet de la fluence conceptuelle sur les RCJ par une fluence de réponse. De plus, la congruence sémantique des images à l'encodage affectait également les prédictions. Les participants prédisaient mieux reconnaître les items appris dans un contexte congruent plutôt qu'incongruent. Puisque les performances mnésiques étaient vraiment affectées par la congruence sémantique, nous avons discuté des possibles facteurs mnésiques en jeu.

**Tableau 8.1** : Récapitulatif des résultats de cette thèse.

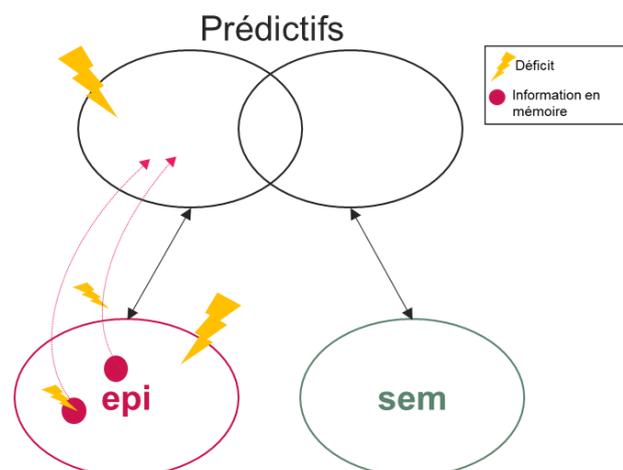
Chapitre	Méthode et population d'intérêt	Jugement	Hypothèses	Confirmation de l'hypothèse ?	Résultats
Ch 4	Méta-analyse Vieillessement sain	eFOK/ sFOK	Le déclin de la sensibilité du eFOK dans le vieillissement découle du déclin en mémoire épisodique	✓	Déclin de la sensibilité du eFOK qui peut venir de facteurs statistiques et mnésiques Pas de preuve d'un déclin du sFOK
Ch 5	Neuropsychologie DS/ MA	ePoR/ sPoR/ eRCJ/ sRCJ	Les personnes DS auraient une moins bonne sensibilité sur tâches sémantiques, surtout pour les PoR	✗	Pas de déficit métacognitif général ou spécifique dans la DS Pas de double dissociation
Ch 6	Psychologie expérimentale Participants jeunes sains	ePoR	Traitement sémantique fournit un cadre aux PoR : moins bonne sensibilité pour prédire la récupération d'informations qui n'ont pas de sens	✓	Meilleure sensibilité métacognitive de PoR lorsque les informations ont du sens Semble venir de meilleures capacités de rappel
Ch 7	Psychologie expérimentale Participants jeunes sains	Exp 7.1 : eRCJ  Exp 7.2 : ePoR	Congruence sémantique utilisée comme indice : confiance plus faible lorsque la cible est traitée de manière moins fluente	≈  ✓	RCJ plus faibles quand distracteur incongruent  PoR plus faibles pour les cibles incongruentes

✓ = hypothèses confirmées ; ✗ = hypothèses non confirmées ; ≈ = hypothèses confirmées en partie

## 8.2. Apport des résultats aux théories de la métacognition

### 8.2.1. Lien entre déficit de premier ordre et sensibilité métacognitive

Nos résultats apportent des preuves à l'hypothèse que des facteurs épisodiques sont impliqués dans les jugements métacognitifs prédictifs faits sur des tâches épisodiques. En effet, les résultats du Chapitre 4 confirment qu'un déclin en ME s'accompagne d'une réduction de la sensibilité métacognitive chez les personnes âgées, sur des tâches épisodiques (voir Figure 8.2). Cependant, quand les performances mnésiques de rappel ou de reconnaissance sont rendues égales de manière artificielle/statistique, les différences de sensibilité liées à l'âge disparaissent. De plus, ce déclin en ME ne s'accompagne pas d'un déclin de sensibilité sur les tâches sémantiques. Ainsi, le chapitre 4 confirme une partie du lien entre mémoire et métacognition : une altération de premier ordre épisodique s'accompagne spécifiquement d'un déficit de second-ordre épisodique sur des tâches de jugements prédictifs (voir Figure 8.2). Nous avons proposé deux raisons possibles à cela. Une première raison est liée aux mesures classiquement utilisées pour évaluer la sensibilité métacognitive (e.g., corrélations gamma,  $d'$  de type-II) puisque la performance de premier ordre est une variable confondue dans le calcul (Galvin et al., 2003; Masson & Rotello, 2009), qui ne donne alors pas une représentation des compétences métacognitives « pures ». Une seconde raison provient du lien intrinsèque entre mémoire et métacognition. Des processus mnésiques épisodiques sont mises en jeu dans une tâche de FOK (e.g., recollection ; Isingrini et al., 2016; Mazancieux et al., 2020; Souchay et al., 2007) et les résultats de la recherche en mémoire peuvent être utilisés comme indice pour les prédictions. Un manque d'accès ou une mauvaise utilisation des informations diminue donc la valeur prédictive des FOK et réduit donc leur sensibilité.

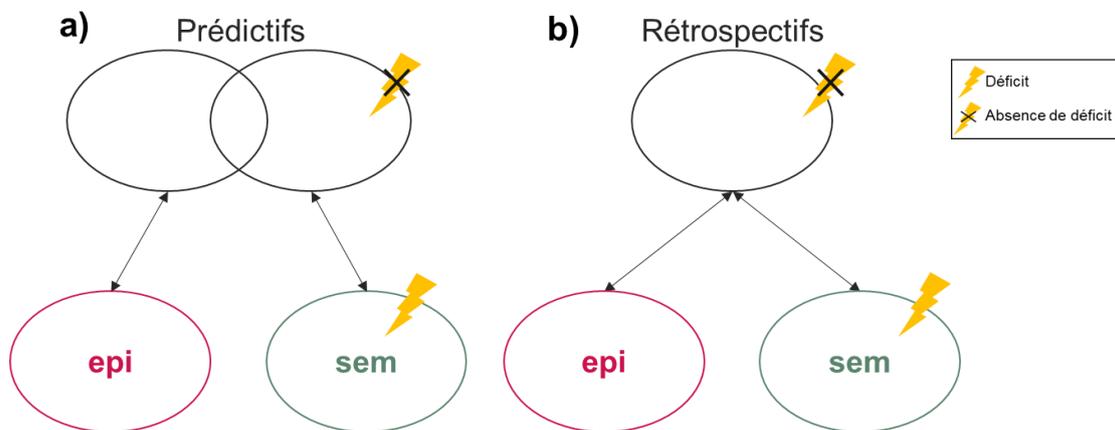


**Figure 8.2 :** Intégration des résultats des 2 méta-analyses du Chapitre 4.

Dans le vieillissement, un déclin de la ME (en-bas) s'accompagne d'un déficit de sensibilité métacognitive (en-haut), qui résulte en partie d'un manque d'accès (i.e., réduction de recollection) ou d'une mauvaise utilisation des informations stockées en ME (points roses) qui guident les jugements. Notons cependant que nous ne considérons pas que ce lien est strict et systématique (e.g., amnésie).

Cette idée de l'implication de facteurs spécifiques à la ME dans les jugements prédictifs comme le FOK, proposée dans certaines théories comme la *memory-constraint hypothesis* (Hertzog et al., 2010), semble donc confirmée par nos méta-analyses.

Cependant, les résultats du Chapitre 5 montrent que ce lien entre déficit de premier ordre et réduction de la sensibilité métacognitive n'est pas retrouvé lorsque le déficit mnésique au premier plan de la pathologie est sémantique. En effet, les personnes atteintes de DS semblent évaluer de manière correcte leurs performances mnésiques, qu'elles soient sémantiques ou épisodiques. Ainsi, ces résultats apportent une preuve supplémentaire de la relation complexe entre mémoire et métacognition : un déficit mnésique n'entraîne pas systématiquement une réduction des capacités métacognitives, même lorsque la forme de mémoire mis en jeu dans la tâche de premier ordre est altérée (voir Figure 8.3). De plus, les résultats obtenus par la méthode des scores de différence sur des questionnaires auto-rapportés montrent que les patients DS sous-estiment leurs difficultés mnésiques dans la vie de tous les jours. Ainsi, d'un point de vue plus clinique, nos résultats montrent la nécessité d'utiliser les paradigmes de métacognition, en complément des mesures plus subjectives et moins contrôlées comme les questionnaires et les entretiens, pour avoir une vue plus détaillée des capacités de monitoring de personnes présentant un trouble de la mémoire (e.g., Cosentino et al., 2011; DeLozier & Davalos, 2016).



**Figure 8.3 :** Intégration des résultats du Chapitre 5. Un déficit mnésique (en-bas) sémantique ne s'accompagne pas nécessairement d'un déficit métacognitif (en-haut), que ce soit pour des jugements prédictifs (a) ou rétrospectifs (b).

### 8.2.2. Influence des facteurs sémantiques dans la métacognition épisodique

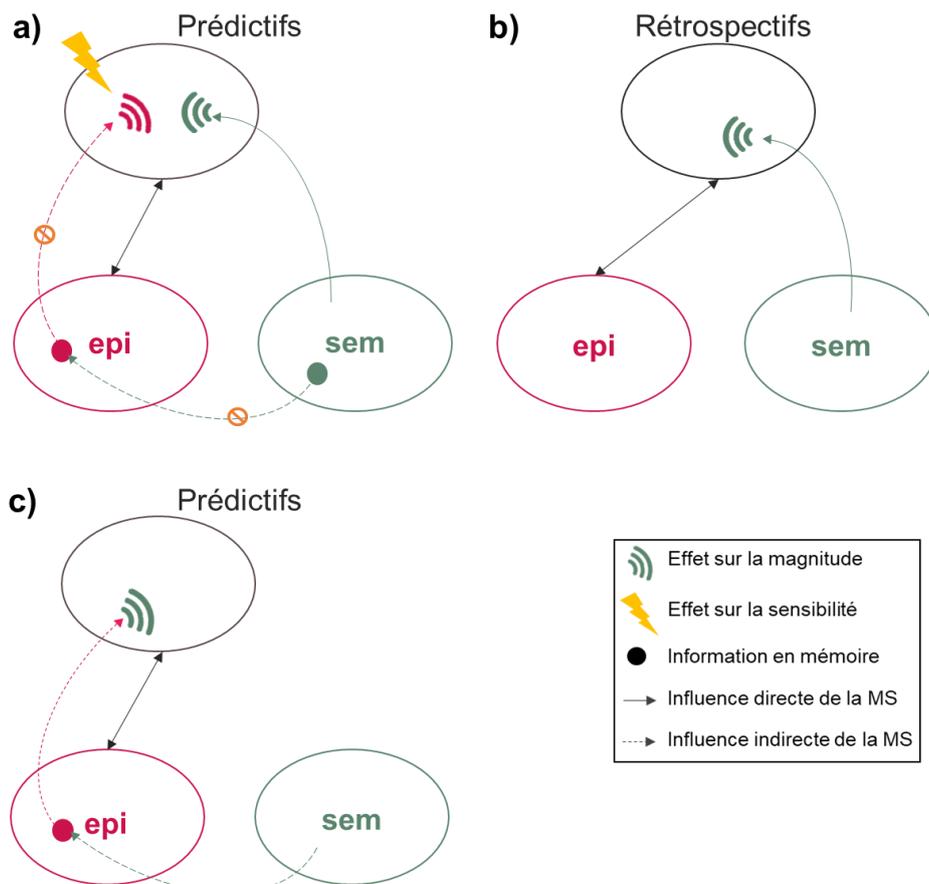
Les résultats du Chapitre 5 suggèrent que la préservation du système sémantique ne semble pas nécessaire aux jugements métacognitifs, qu'ils soient prédictifs ou rétrospectifs

(voir Figure 8.3). Cependant, les Chapitre 6 et 7 montrent que chez des sujets sains, le traitement sémantique des informations à retenir influencent les jugements métacognitifs. Nous proposons deux types d'influence de ces facteurs sémantiques : directs et indirects, via des facteurs épisodiques (voir Figure 8.4).

Tout d'abord, les expériences du Chapitre 6 (voir Figure 8.4a) ainsi que l'expérience 7.2 (voir Figure 8.4c) fournissent des exemples d'influence directe de la MS sur les jugements métacognitifs épisodiques. En effet, dans le Chapitre 6, les participants prédisaient moins de reconnaissance lorsqu'un indice n'ayant pas de sens était présenté, même lorsque la cible n'était pas rappelée. Ainsi, il semble qu'un indice métacognitif intrinsèque, lié à l'item indice des paires, aient été utilisé pour inférer directement une récupération future de la cible. Cet indice était le fait de ne pas avoir de représentation sémantique des items présentés au moment de la prédiction. Puisque les cibles qui ont été apprises avec un indice n'ayant pas de sens étaient réellement moins bien reconnues, d'autres travaux sont nécessaires afin de comprendre si cet indice métacognitif intrinsèque influence la magnitude des prédictions via un mécanisme basé sur l'expérience (e.g., sentiment de familiarité, force mnésique plus faible) ou sur des théories (e.g., croyance que la mémoire des cibles indicées par un pseudo-mot sera plus faible). De plus, dans l'expérience 7.1 (voir Figure 8.4b), les participants étaient moins certains de l'exactitude de leurs réponses dans les cas où les distracteurs étaient dans un contexte incongruent avec leurs connaissances. Leur mémoire n'étant pas influencée par la congruence des cibles et des distracteurs, il semble que la congruence influençait directement la magnitude des jugements rétrospectifs, et non par un facteur mnésique. Ainsi, dans ce cas, la MS, mise en jeu par l'influence de la congruence sur le traitement des items, était utilisée pour inférer l'exactitude des réponses. Cette influence semblait venir d'un ralentissement des réponses dans le cas où les distracteurs étaient incongruents que nous interprétons comme reflétant une difficulté de traitement, et donc de facteurs basés sur l'expérience.

D'un autre côté, les expériences 6.4 et 7.2 suggèrent également une influence plus indirecte de la MS sur les jugements métacognitifs prédictifs et leur sensibilité, via une influence sur la ME (voir Figure 8.4a et Figure 8.4c). En effet, la MS influence le traitement des informations à apprendre, leur encodage et leur récupération en ME. Comme nous en avons discuté dans la partie précédente, nous proposons que ces effets sur la récupération en ME influencent ensuite les prédictions faites en mémoire épisodique. Notamment, dans l'expérience 6.4, empêcher le traitement sémantique des informations à apprendre en utilisant des pseudo-mots, limitait les capacités à récupérer les informations en mémoire au moment de faire la prédiction (voir Figure 8.4a). Lorsqu'ils connaissaient l'indice de la paire présenté, les participants prédisaient se souvenir de moins de cibles n'ayant pas de représentation sémantique que de cibles ayant une signification. Cela était expliqué par le fait que lorsqu'ils

ne pouvaient pas utiliser leurs représentations sémantiques et créer un lien sémantique entre les deux items de la paire (paires avec au moins un pseudo-mots), ils n'arrivaient à rappeler la cible au moment de faire leurs PoR. De ce fait, les cibles étant alors moins accessibles, les prédictions reflétaient moins la performance de premier ordre et les participants avaient alors une moins bonne sensibilité métacognitive. De même, dans l'expérience 7.2, les participants prédisaient pouvoir plus se souvenir des items ayant été appris dans un contexte congruent avec leurs connaissances plutôt qu'incongruents. Nous proposons que cela est dû au fait que d'avoir des connaissances sur l'endroit où les items sont généralement rencontrés dans la vie de tous les jours influence le traitement de ces items et leur mémorisation (Ortiz-Tudela et al., 2017), puisque les participants ont mieux reconnu les items congruents qu'incongruents. Ensuite, au moment où ils font les PoR, les participants auraient plus accès aux cibles apprises dans un contexte congruent, et ajusteraient donc leurs prédictions en conséquence (i.e., magnitude plus élevée). Cependant, des études supplémentaires sont nécessaires afin de vérifier que des facteurs mnésiques épisodiques sont bien à l'origine du changement de la magnitude des jugements dans cette expérience 7.2.



**Figure 8.4 :** Intégration des résultats des Chapitres 6 (a) et 7 (b et c). La MS peut jouer sur la magnitude des jugements prédictifs (a et c) et rétrospectifs (b) à travers des influences que nous supposons comme directes (trait plein), ou indirectes via la ME (en pointillés).

### 8.3. Limites et perspectives

Plusieurs limites à nos études sont à considérer. Certaines découlent des paradigmes utilisés alors que d'autres concernent des questionnements qui restent en suspens. Tout d'abord, Vuorre & Metcalfe (2021) proposent que bien que de précédents travaux aient suggéré que les mesures de sensibilité métacognitive ne dépendent pas de la performance de premier ordre (e.g., Nelson, 1984), cette conclusion dépend des paradigmes utilisés. En particulier, un élément important pouvant influencer l'interdépendance des mesures de premier et de second ordre est le type de tâche de premier ordre : rappel libre, paradigme ancien/nouveau (i.e., *old/new* en anglais), ou reconnaissance en choix multiples. Notamment, ces auteurs ont montré que plus il y a d'alternatives proposées dans une tâche de reconnaissance en choix multiples, meilleure est la sensibilité métacognitive. Cela serait dû au fait que la possibilité de choisir la bonne réponse au hasard (i.e., « *guessing* ») diminue avec l'augmentation du nombre de distracteurs (Schwartz & Metcalfe, 1994). Cependant, dans toutes les études présentées dans cette thèse, la tâche de premier ordre était un test de reconnaissance en choix forcé à 2 alternatives. Dans ces tâches, une personne qui répond au hasard devrait indiquer une confiance faible en sa réponse. Or, dans un paradigme avec seulement deux alternatives, la probabilité de répondre correctement mais au hasard est dans ce cas élevée (0.5) et il y a une chance sur deux pour que l'essai soit alors considéré comme une omission de type-2 (i.e., reconnaissance correcte mais faible jugement), ce qui réduit la sensibilité métacognitive. Au moment des analyses, il est impossible de distinguer les essais corrects répondus au hasard des essais corrects répondus volontairement. Ce choix de paradigme était contraint par les prérequis à l'utilisation de mesures d'efficacité métacognitive (Maniscalco & Lau, 2014), que nous avons initialement prévue de mesurer sur l'ensemble de nos études. Cependant, dans les tâches avec peu d'essais, comme c'était notamment le cas dans le Chapitre 5 (20 essais par tâche) et 6 (30 essais par conditions), le poids de ces essais réussis malgré une réponse au hasard, identifié comme tels par les participants (i.e., magnitude faible) est d'autant plus important dans la mesure de la sensibilité.

Une autre limite sous-tend l'utilisation de paradigmes en choix forcé à 2 alternatives : il est possible de trouver la réponse correcte par familiarité avec ou recollection de la cible, ou par rejet du distracteur (e.g., « je n'ai pas vu cet item » ; i.e., heuristique de distinctivité ; Schacter et al., 1999). Dans le cas des paradigmes de prédictions, les participants ne voient pas encore les distracteurs au moment de faire leur jugement, ainsi ils peuvent prédire ne pas reconnaître la cible puis ensuite sélectionner la bonne cible par rejet du distracteur. La prédiction est alors considérée comme une omission puisque la réponse était bien correcte, et diminue donc la sensibilité, bien que conceptuellement le participant ait correctement évalué

le fait de ne pas connaître la cible. Une manière d'éliminer la probabilité d'utiliser ce rejet par distinctivité est d'augmenter le nombre de distracteurs.

On peut alors se demander si la réduction de la sensibilité observée dans les conditions avec des pseudo-mots dans le Chapitre 6 vient, du moins en partie, d'une plus grande utilisation des réponses au hasard dans ces conditions ou de rejet du distracteur. En effet, il est possible que dans ces conditions, en découvrant certains pseudo-mots distracteurs, les sensations ressenties lors de leur découverte (« ce pseudo-mot me fait penser au nom de mon chien) soient utilisées pour inférer que ces items n'aient pas été vus avant (« je m'en serais souvenu »). De plus, dans les expériences des chapitres 5 et 6, des différences de performances de premier ordre étaient observées, même si elles ne s'accompagnaient pas toujours de différences au niveau des performances de second ordre. L'interdépendance entre performance de premier et de second ordre étant d'autant plus forte dans des paradigmes avec 2 alternatives, il serait intéressant de voir si les résultats obtenus seraient identiques avec des paradigmes avec un plus grand nombre de distracteurs.

Ensuite, dans l'expérience du Chapitre 5 réalisée auprès de personnes atteintes de DS, sur les bases des travaux suggérant un lien statistique entre difficulté de la tâche et mesures de métacognition (Galvin et al., 2003; Masson & Rotello, 2009), plusieurs versions des tâches ont été utilisées. La première raison était d'éviter des performances plafond chez les participants sains et planchers chez les patients, afin qu'il soit pertinent d'utiliser les mesures de monitoring. En effet, pour rappel, nous nous intéressons à la sensibilité métacognitive qui est la capacité à discriminer les essais corrects des essais incorrects (Schraw, 2009). Il n'aurait donc pas été pertinent de l'étudier si tous les essais étaient réussis ou si la tâche était tellement difficile que les participants répondaient au hasard sur tous les essais. La seconde raison était de réduire les différences de performances entre les patients et les contrôles, afin de mesurer les capacités métacognitives plus « pures » (i.e., non contaminée par la performance de premier ordre). L'utilisation des différentes versions sur les tâches de PoR sémantique a en réalité entraîné des performances de premier ordre plus élevées chez les personnes DS que chez les contrôles. Ainsi, si les résultats montrent qu'il n'y a pas de déficit métacognitif « pur » dans cette pathologie, on peut se demander ce qui se serait passé si les performances avaient été égales, ou si une même tâche avait été utilisée pour les différents groupes. Il pourrait donc être pertinent d'augmenter le nombre d'items et d'intégrer à l'intérieur à la fois des essais simples et des essais difficiles (sur la base de l'étude pilote). Cela permettrait d'avoir des essais sur lesquels les personnes DS devraient pouvoir répondre et d'autres plus compliqués, auxquels les participants sains auraient plus de mal à répondre. La conséquence serait normalement une plus faible performance de premier ordre chez les personnes DS que chez les contrôles, qui contaminerait les mesures de sensibilité.

Dans la vie de tous les jours, les personnes DS vivent avec leur déficit mnésique : elles ont notamment du mal à identifier les personnes et à reconnaître certains objets (Hodges & Patterson, 2007). De plus, comme nous en avons déjà discuté, un déficit mnésique n'entraîne pas nécessairement un déficit métacognitif. D'un point de vue plus clinique, il est donc pertinent de se poser la question des capacités métacognitives sur des tâches identiques à celles des participants sains, et donc dans le cas où elles devraient présenter un déficit sémantique, même si les mesures sont contaminées par la performance de premier ordre.

Ces remarques reflètent donc un dilemme dans le choix des paradigmes à utiliser lorsque l'on cherche à évaluer la métacognition dans des pathologies de la mémoire. D'un point de vue fondamental/théorique, il est nécessaire de contrôler la performance de premier ordre pour évaluer les capacités métacognitives « pures » et mieux comprendre les processus sous-tendant les jugements. D'un point de vue clinique, il est plus pertinent de mettre les participants DS dans des situations qui reflètent leurs difficultés dans la vie de tous les jours, pour voir s'ils arrivent à estimer l'exactitude de leurs connaissances et souvenirs. Pour cette raison mais aussi pour comprendre la relation entre mémoire et métacognition, il semble important d'évaluer les capacités métacognitives sur des tâches sur lesquelles ils présentent un déclin des performances mnésiques. Les approches fondamentales et cliniques apportent donc des éléments complémentaires : d'un côté pour la compréhension du fonctionnement du processus de monitoring ; de l'autre, pour la compréhension de l'évaluation que les patients font de leurs capacités mnésiques.

D'un autre côté, d'autres études pourraient être conduites pour compléter les résultats des expériences du Chapitre 6 concernant le rôle des traitements et des représentations sémantiques. Plusieurs aspects du traitement sémantique du matériel utilisé pourraient intervenir comme l'imagerie ou la récupération d'informations partielles sémantiques. Afin de déterminer si les effets observés sont une conséquence de la possibilité d'utiliser l'imagerie mentale, la même expérience pourrait être réalisée à l'aide de pseudo-mots, de mots abstraits, et de mots concrets (Paivio et al., 1994). De plus, il pourrait être intéressant de réaliser le même type d'expérience auprès de patients DS avec des mots ou images connues et d'autres inconnues. Ainsi, en classifiant par personne les items connus et non reconnus (i.e., représentation sémantique supposée altérée), on pourrait étudier l'apport des représentations sémantiques intactes pour les prédictions de reconnaissance.

Par ailleurs, plusieurs autres questions restent en suspens. Si nous avons montré que les connaissances sémantiques influencent la magnitude des jugements rétrospectifs et prédictifs, les processus sous-jacents ont besoin d'être davantage investigués. Dans le Chapitre 6, les participants prédisaient moins reconnaître les cibles apprises avec des indices

n'ayant pas de sens, et reconnaissaient vraiment moins ces cibles que les cibles indicées par des mots réels. Il est donc difficile de distinguer si la magnitude était influencée par des facteurs mnésiques et le fait que ces cibles étaient moins accessibles au moment de la prédiction, si les participants utilisaient la familiarité avec l'indice vu au moment de leur prédiction, ou s'ils inféraient leurs prédictions d'une croyance que les cibles étaient moins bien retenues lorsqu'elles étaient précédées par un mot sans signification. De plus, dans le Chapitre 7, les connaissances étaient utilisées comme indices mais ces indices étaient-ils basés sur des expériences et sensations ressenties vis-à-vis de la récupération en mémoire ? Ou plutôt basé sur des théories sur l'effet de la congruence sur leur mémoire ? Si les résultats de l'expérience 7.1 suggèrent que l'effet de la congruence sur les jugements rétrospectifs découle de facteurs basés sur l'expérience, comme la fluence de récupération, l'expérience 7.2 ne permet pas de tirer de conclusions sur l'origine de l'effet de congruence sur les prédictions. Il pourrait être intéressant de refaire la même expérience en inversant le pattern de performance mnésique, comme cela était fait dans l'expérience de Carroll et al. (1997, voir 50). Le principe serait de faire apprendre plusieurs fois les images aux participants, jusqu'à ce qu'ils soient capables de rappeler les cibles sur présentation des scènes, mais avec un critère de rappel plus important pour les paires incongruentes que congruentes. La dissociation entre congruence et degré d'apprentissage permettrait de voir si les participants utilisent des croyances (i.e., ils se basent sur la congruence) ou plutôt l'expérience mnésique (i.e., ils se basent sur le degré d'apprentissage).

De plus, de nombreuses questions persistent pour comprendre la généralité des jugements métacognitifs. Par exemple, il pourrait également être intéressant de comparer dans un même paradigme l'effet de la congruence sémantique sur les prédictions et les jugements rétrospectifs. Cela permettrait de voir plus spécifiquement si les connaissances générales influencent de la même manière et au même degré les jugements prédictifs et rétrospectifs.

Enfin, un aspect de la relation entre mémoire et métacognition nécessiterait d'être plus profondément étudié : quels sont les indices utilisés et mécanismes sous-tendant les jugements métacognitifs faits sur des tâches de mémoire sémantique ? Y a-t-il des mécanismes ou indices utilisés exclusivement dans les tâches de métacognition sémantique et non épisodiques ? De plus, un aspect de la mémoire sémantique est son caractère multimodal (e.g., Binder & Desai, 2011; Farah et al., 1989; Lambon Ralph & Patterson, 2008; Rogers et al., 2004). On peut accéder à un concept par les modalités visuelle (e.g., image d'une fleur), auditive (e.g., mot « fleur »), ou encore tactile (e.g., toucher une fleur) et olfactive (e.g., sentir l'odeur d'une fleur). La majorité des études sur la métacognition sémantique ont été réalisées avec du matériel verbal (e.g., questions de culture générale) ou visuel (e.g.,

reconnaissance de visages, voir [Annexe A](#) pour des exemples dans le vieillissement). Quelques études ont porté sur la métacognition dans l'olfaction (Jönsson et al., 2005; Jönsson & Olsson, 2003). Il serait donc intéressant de voir si les capacités métacognitives des personnes diffèrent d'une modalité d'entrée à une autre, à la fois chez des participants sains, et chez des personnes avec un trouble affectant supposément toutes les modalités sensorielles comme la DS. On pourrait également évaluer si le potentiel effet de la modalité de présentation est le même dans la métacognition épisodique.

# CONCLUSION DE LA THÈSE

---

Nos travaux apportent des preuves supplémentaires au rôle de la mémoire dans les capacités métacognitives. En particulier, les compétences épisodiques comme la récupération en mémoire sont impliquées pour les jugements épisodiques, notamment prédictifs. Ainsi, un déclin des capacités de premier ordre épisodiques peut être à l'origine d'une réduction de la sensibilité de ces jugements prédictifs épisodiques. Ce résultat semble aller dans les sens des études qui suggèrent une spécificité des capacités métacognitives lorsque des jugements prédictifs sont étudiés. Cependant, cette relation entre déficit mnésique et métacognition est complexe. Par exemple, un déficit sémantique n'entraîne pas nécessairement un déclin de la sensibilité métacognitive, même sur des tâches de mémoire sémantique. De plus, nos résultats montrent que des informations sémantiques sont également mises en jeu dans les jugements métacognitifs épisodiques. Nous avons discuté des potentielles influences directes et indirectes des connaissances sémantiques sur les performances métacognitives épisodiques. Globalement, nous suggérons que nos travaux apportent des éléments pour mieux comprendre comment le monitoring fonctionne, et soulignent l'importance de la mémoire épisodique et de la mémoire sémantique dans ce processus.

## Annexe A. Chapitre 4 : Résumé des études incluses dans la méta-analyse globale.

OA = Older adults. YA = Young adults. XAFC = X-alternative-forced-choice task. \* This 30-minutes condition was not included in the meta-analysis as data were missing.

Authors (year)	Sample size		Task	Material	Recognition type	FOK scale	Number of trials	Attempt of matched performance
	YA	OA						
Allen-Burge & Storandt (2000) <sup>43</sup> – Exp 1	45	45	Semantic	Rare-word definitions	2AFC	7-point	50	No
Butterfield et al. (1988) <sup>14</sup> – Exp 2	54	36	Semantic	General-knowledge questions	7AFC	Yes/no	12	No
Douchemane et al. (2007) <sup>15</sup>	18	18	Semantic	Definitions of words	5AFC	Yes/no	40	No
			Episodic	Pairs of words	5AFC	Yes/no	40	No
Eakin & Hertzog (2012) <sup>22</sup>	51	42	Episodic	Pairs of words	5AFC	0-100%	44 (22+22)	No
Eakin et al. (2014) <sup>21</sup>	50	56	Semantic	Famous faces' names	3AFC	0-100%	30	No
			Episodic	Non-famous faces with names (rearranged from famous names)	3AFC	0-100%	30	No
Hertzog et al. (2010) <sup>24</sup>	54	54	Episodic	Pairs of words	4AFC	0-100%	20	Manipulation of the delay between encoding and test. Larger delay for the younger group (7 days). Two OA groups: one with a delay of 48 hours and the other with a delay of 30 minutes*.
Maclaverty & Hertzog (2009) <sup>42</sup>	206	200	Episodic	Pairs of words	4AFC	25%-100%	36	No
Marquie & Huet (2000) <sup>7</sup>	22	22	Semantic	General-knowledge + computer-related questions	4AFC	5-point	138 (69+69)	No
Morson et al. (2015) <sup>16</sup>	35	16	Semantic	General-knowledge questions	4AFC	Yes/no	60	No
			Episodic	Answers to general-knowledge questions (unrecognized in semantic memory)	4AFC	Yes/no	60 minus items with correct recognition on the sFOK task	No
Perrottin et al. (2006) <sup>29</sup>	40	62	Episodic	Pairs of words	5AFC	Yes/no	40	No
Sueher et al. (2013) <sup>40</sup>	20	60	Episodic	Pairs of words	5AFC	6-point (0-100%)	0-60	Manipulation of attention. These YA groups: one group with divided attention at encoding, the other with divided attention for FOK judgment. The third group was a control group with full attention.
Sueher et al. (2015) <sup>27</sup>	59	61	Episodic	Pairs of words	5AFC	0-100%	60	No
Souchay & Isingrini (2012) <sup>26</sup>	16	36	Episodic	Pairs of words	Old/new	Yes/no	40	No
Souchay et al. (2000) <sup>25</sup>	20	41	Episodic	Pairs of words	Yes/no	Yes/no	36	No
Souchay et al. (2002) <sup>23</sup>	16	16	Episodic	Pairs of words	Old/new	Yes/no	20	No
Souchay et al. (2007) <sup>17</sup> – Exp 1	20	40	Semantic	General-knowledge questions	5AFC	Yes/no	40	No
			Episodic	Pairs of words	5AFC	Yes/no	40	No
Souchay et al. (2007) <sup>17</sup> – Exp 2	20	36	Episodic	Pairs of words	5AFC	Yes/no	40	No
Thomas et al. (2011) – Exp 1	42	42	Episodic	Pairs of words	6AFC	17-100%	36	Manipulation of the presentation time at encoding (500ms for YA and 5s for OA). Participants could only move to the FOK phase if they had at least 33% of correct recall.
Thomas et al. (2011) <sup>41</sup> – Exp 2	22	20	Episodic	Pairs of words	6AFC	17-100%	36	Same as Experiment 1.
Thomas et al. (2011) <sup>41</sup> – Exp 3	24	15	Episodic	Pairs of words	6AFC	17-100%	36	Same as Experiments 1 and 2.

**Annexe B. Chapitre 5 : Liste des mots utilisé pour la tâche de PoR sémantique.**

Niveau facile	Niveau intermédiaire	Niveau difficile
accordéon agrafeuse ancre arrosoir ceinture concombre crevette échelle guitare horloge kayak louche oignon passoire raisin rasoir seringue tablier télescope valise	caribou iguane langouste trèfle arbalète stéthoscope opossum tipi écrou tétine calèche brouette trampoline mitaine sécateur jarre métronome carillon enclume dague	alène bédane daman daubière fardier gabare gimblette grelinette gyroscope licol mousquet otoscope papinette piccolo sapote serfouette stroboscope tamarin thérémine wombat

**Annexe C. Chapitre 5 : Tableau récapitulatif des moyennes (écart-types) pour chaque tâche et groupe.**

Tâche			Contrôles	DS	MA		
PoR	Épisodique	%RC	0.90 (0.08)	0.80 (0.11)	0.69 (0.24)		
		Magnitude	Tous	0.71 (0.11)	0.68 (0.14)	0.63 (0.09)	
			<i>Essais corrects</i>	0.71 (0.11)	0.68 (0.14)	0.63 (0.10)	
			<i>Essais incorrects</i>	0.67 (0.13)	0.65 (0.14)	0.62 (0.09)	
		AUROC2	0.50 (0.09)	0.50 (0.10)	0.45 (0.07)		
	Sémantique	%RC	0.72 (0.09)	0.83 (0.12)	0.74 (0.02)		
		Magnitude	Tous	0.71 (0.09)	0.79 (0.13)	0.61 (0.07)	
			<i>Essais corrects</i>	0.74 (0.11)	0.81 (0.13)	0.63 (0.09)	
			<i>Essais incorrects</i>	0.63 (0.08)	0.69 (0.17)	0.57 (0.06)	
		AUROC2	0.60 (0.12)	0.61 (0.11)	0.52 (0.10)		
	RCJ	Épisodique	%RC	0.87 (0.09)	0.66 (0.16)	0.75 (0.15)	
			Magnitude	Tous	0.88 (0.11)	0.83 (0.13)	0.76 (0.09)
				<i>Essais corrects</i>	0.89 (0.11)	0.83 (0.12)	0.76 (0.09)
				<i>Essais incorrects</i>	0.83 (0.13)	0.82 (0.15)	0.76 (0.10)
AUROC2			0.69 (0.11)	0.56 (0.11)	0.52 (0.04)		
Sémantique		%RC	0.77 (0.15)	0.68 (0.16)	0.65 (0.15)		
		Magnitude	Tous	0.82 (0.11)	0.79 (0.14)	0.76 (0.09)	
			<i>Essais corrects</i>	0.83 (0.11)	0.81 (0.15)	0.78 (0.09)	
			<i>Essais incorrects</i>	0.74 (0.14)	0.73 (0.14)	0.73 (0.11)	
		AUROC2	0.62 (0.12)	0.62 (0.12)	0.57 (0.15)		

%RC = proportion de reconnaissance correcte

## Annexe D. Chapitre 6 : Exemples de paires utilisées pour un bloc.

Indice	Cible	Distracteur	Type de paire
narine	toilette	tondeuse	
fromage	carnet	anneau	
lavande	corbeille	canon	
raquette	aiguille	moulin	
perruque	rectangle	tonneau	
renard	affiche	costume	
bouilloire	dragon	bougie	
rasoir	poteau	torchon	W-W
classeur	flocon	vaisselle	
framboise	marteau	lanterne	
crayon	portail	berceau	
licence	hibou	piano	
sabot	fourrure	moustache	
sonnette	boudin	casquette	
ardoise	barrage	ficelle	
oignon	couraille	linduque	
ciseaux	pelume	ropie	
bocal	voinon	bantime	
pendule	vompis	camose	
tambour	latorne	sopprie	
concombre	crosave	arsage	
cactus	larnit	pongrome	
serviette	omin	satusse	W-PW
maillot	calpide	amaisse	
raisin	setteau	phalson	
orteil	clasion	otoule	
charette	aprolle	chonquet	
grenier	misancle	daguin	
bijou	camprette	allime	
bassin	fulacre	calisme	
culane	toupie	tableau	
darnuse	limace	olive	
abouce	royaume	balance	
ongade	colonne	album	
grinou	manoir	avion	
transsette	bouton	tapis	
palin	usine	marin	
budale	piscine	valise	PW-W
cituse	drapeau	crevette	
haril	pigeon	rocher	
taglette	triangle	pinceau	
bortige	chariot	balai	
roumisme	colline	farine	
ventelle	gravier	coussin	
bocue	cabane	semelle	
trencarde	siguel	baline	
glonin	proumase	mautiche	
tescade	torquet	braniale	
farlonne	socat	ganssion	
brouman	coginte	pocront	
dauffrate	counelle	rureune	
trolice	mutrole	bicret	
splindron	bulas	taudige	PW-PW
penguët	carand	convisse	
houdis	leugie	sepot	
couliade	spairint	poitire	
adue	rattie	bonal	
flaumet	reflon	bloumise	
gaucule	suglet	tramoint	
bulet	achave	pironne	

**Annexe E. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.1.**

		W targets	PW targets	M
<b>Memory performance</b>				
%CR	W cue	0.78 (0.11)	0.72 (0.10)	<b>0.75 (0.11)</b>
	PW cue	0.70 (0.11)	0.66 (0.12)	<b>0.68 (0.11)</b>
	<b>M</b>	<b>0.74 (0.12)</b>	<b>0.69 (0.11)</b>	
d'	W cue	1.55 (0.72)	1.19 (0.63)	<b>1.37 (0.70)</b>
	PW cue	1.06 (0.70)	0.86 (0.64)	<b>0.96 (0.67)</b>
	<b>M</b>	<b>1.31 (0.75)</b>	<b>1.02 (0.65)</b>	
<b>Metacognitive performance</b>				
Magnitude (transformed values)	W cue	0.79 (0.09)	0.75 (0.08)	<b>0.77 (0.09)</b>
	PW cue	0.66 (0.08)	0.66 (0.08)	<b>0.66 (0.08)</b>
	<b>M</b>	<b>0.72 (0.11)</b>	<b>0.70 (0.09)</b>	
Gamma correlations	W cue	0.41 (0.40) <sup>***</sup>	0.08 (0.35)	<b>0.25 (0.41)</b>
	PW cue	0.21 (0.34) <sup>***</sup>	0.15 (0.38) <sup>**</sup>	<b>0.18 (0.36)</b>
	<b>M</b>	<b>0.31 (0.38)</b>	<b>0.12 (0.36)</b>	
AUROC2	W cue	0.62 (0.11) <sup>***</sup>	0.52 (0.09)	<b>0.57 (0.11)</b>
	PW cue	0.53 (0.07) <sup>**</sup>	0.52 (0.09)	<b>0.52 (0.08)</b>
	<b>M</b>	<b>0.58 (0.11)</b>	<b>0.52 (0.09)</b>	

Standard deviations are provided in parentheses. %CR = proportion of correct recognition; PW = pseudoword; W = word. Stars indicate significant differences from 0 for gamma correlations and from 0.5 for AUROC2 (\* :  $p < .01$ ; \*\* :  $p < .001$ ).

## Annexe F. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.2.

Standard deviations are provided in parentheses. %CR = proportion of correct recognition; PW = pseudoword; W = word. Stars

	Group		W targets	PW targets	M
<b>Memory performance</b>					
%CR	Control	W cue	0.77 (0.12)	0.73 (0.10)	<b>0.75 (0.11)</b>
		PW cue	0.71 (0.12)	0.67 (0.10)	<b>0.69 (0.11)</b>
		<b>M</b>	<b>0.74 (0.13)</b>	<b>0.70 (0.11)</b>	<i>0.72 (0.12)</i>
	TM	W cue	0.71 (0.10)	0.75 (0.08)	<b>0.73 (0.09)</b>
		PW cue	0.67 (0.09)	0.69 (0.08)	<b>0.68 (0.09)</b>
		<b>M</b>	<b>0.69 (0.10)</b>	<b>0.72 (0.09)</b>	<i>0.71 (0.09)</i>
d'	Control	W cue	1.52 (0.78)	1.22 (0.62)	<b>1.37 (0.72)</b>
		PW cue	1.17 (0.75)	0.96 (0.59)	<b>1.07 (0.68)</b>
		<b>M</b>	<b>1.34 (0.78)</b>	<b>1.09 (0.61)</b>	<i>1.22 (0.56)</i>
	TM	W cue	1.13 (0.66)	1.34 (0.50)	<b>1.23 (0.59)</b>
		PW cue	0.88 (0.49)	0.99 (0.46)	<b>0.94 (0.48)</b>
		<b>M</b>	<b>1.00 (0.59)</b>	<b>1.17 (0.51)</b>	<i>1.09 (0.71)</i>
<b>Metacognitive performance</b>					
Magnitude (transformed values)	Control	W cue	0.75 (0.09)	0.71 (0.09)	<b>0.73 (0.09)</b>
		PW cue	0.65 (0.09)	0.64 (0.09)	<b>0.65 (0.09)</b>
		<b>M</b>	<b>0.70 (0.10)</b>	<b>0.68 (0.09)</b>	<i>0.69 (0.10)</i>
	TM	W cue	0.71 (0.07)	0.69 (0.07)	<b>0.70 (0.07)</b>
		PW cue	0.64 (0.07)	0.64 (0.07)	<b>0.64 (0.07)</b>
		<b>M</b>	<b>0.68 (0.08)</b>	<b>0.66 (0.07)</b>	<i>0.67 (0.08)</i>
Gamma correlation	Control	W cue	0.49 (0.36) <sup>***</sup>	0.24 (0.38) <sup>***</sup>	<b>0.37 (0.38)</b>
		PW cue	0.21 (0.47) <sup>**</sup>	0.22 (0.46) <sup>**</sup>	<b>0.22 (0.46)</b>
		<b>M</b>	<b>0.35 (0.43)</b>	<b>0.23 (0.42)</b>	<i>0.29 (0.43)</i>
	TM	W cue	0.35 (0.37) <sup>***</sup>	0.13 (0.40) <sup>*</sup>	<b>0.24 (0.40)</b>
		PW cue	0.21 (0.46) <sup>**</sup>	0.14 (0.47)	<b>0.17 (0.46)</b>
		<b>M</b>	<b>0.28 (0.42)</b>	<b>0.14 (0.44)</b>	<i>0.21 (0.43)</i>
AUROC2	Control	W cue	0.63 (0.10) <sup>***</sup>	0.55 (0.09) <sup>**</sup>	<b>0.59 (0.10)</b>
		PW cue	0.51 (0.08)	0.52 (0.08)	<b>0.52 (0.07)</b>
		<b>M</b>	<b>0.57 (0.11)</b>	<b>0.53 (0.08)</b>	<i>0.55 (0.10)</i>
	TM	W cue	0.59 (0.10) <sup>***</sup>	0.51 (0.09)	<b>0.55 (0.10)</b>
		PW cue	0.52 (0.08)	0.50 (0.08)	<b>0.51 (0.08)</b>
		<b>M</b>	<b>0.55 (0.10)</b>	<b>0.51 (0.09)</b>	<i>0.53 (0.09)</i>

indicate significant differences from 0 for bias and gamma correlations and from 0.5 for AUROC2 (\*\* : p < .01; \*\*\* : p < .001).

**Annexe G. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.3.**

	Group		W targets	PW targets	M
<b>Memory performance</b>					
%CR	Control	W cue	0.74 (0.12)	0.72 (0.09)	<b>0.73 (0.11)</b>
		PW cue	0.69 (0.11)	0.68 (0.11)	<b>0.68 (0.11)</b>
		<b>M</b>	<b>0.72 (0.12)</b>	<b>0.70 (0.10)</b>	<i>0.71 (0.11)</i>
	Familiarised	W cue	0.74 (0.11)	0.78 (0.10)	<b>0.76 (0.11)</b>
		PW cue	0.70 (0.10)	0.73 (0.12)	<b>0.71 (0.11)</b>
		<b>M</b>	<b>0.72 (0.11)</b>	<b>0.76 (0.11)</b>	<i>0.74 (0.11)</i>
d'	Control	W cue	1.34 (0.78)	1.17 (0.54)	<b>1.25 (0.68)</b>
		PW cue	0.98 (0.63)	0.95 (0.61)	<b>0.97 (0.62)</b>
		<b>M</b>	<b>1.16 (0.73)</b>	<b>1.06 (0.58)</b>	<i>1.11 (0.66)</i>
	Familiarised	W cue	1.32 (0.69)	1.60 (0.68)	<b>1.46 (0.70)</b>
		PW cue	1.06 (0.64)	1.25 (0.73)	<b>1.16 (0.69)</b>
		<b>M</b>	<b>1.19 (0.67)</b>	<b>1.42 (0.72)</b>	<i>1.31 (0.71)</i>
<b>Metacognitive performance</b>					
Magnitude (transformed values)	Control	W cue	0.72 (0.10)	0.69 (0.08)	<b>0.70 (0.09)</b>
		PW cue	0.64 (0.06)	0.64 (0.06)	<b>0.64 (0.06)</b>
		<b>M</b>	<b>0.68 (0.09)</b>	<b>0.66 (0.08)</b>	<i>0.67 (0.08)</i>
	Familiarised	W cue	0.73 (0.09)	0.70 (0.08)	<b>0.71 (0.09)</b>
		PW cue	0.65 (0.08)	0.65 (0.08)	<b>0.65 (0.08)</b>
		<b>M</b>	<b>0.69 (0.09)</b>	<b>0.67 (0.09)</b>	<i>0.68 (0.09)</i>
Gamma correlation	Control	W cue	0.42 (0.39)***	0.13 (0.40)*	<b>0.27 (0.42)</b>
		PW cue	0.26 (0.36)***	0.16 (0.29)***	<b>0.21 (0.33)</b>
		<b>M</b>	<b>0.34 (0.38)</b>	<b>0.15 (0.35)</b>	<i>0.24 (0.38)</i>
	Familiarised	W cue	0.43 (0.39)***	0.14 (0.46)*	<b>0.28 (0.45)</b>
		PW cue	0.25 (0.35)***	0.21 (0.49)**	<b>0.23 (0.43)</b>
		<b>M</b>	<b>0.34 (0.38)</b>	<b>0.18 (0.47)</b>	<i>0.26 (0.44)</i>
AUROC2	Control	W cue	0.61 (0.11)***	0.52 (0.09)	<b>0.56 (0.11)</b>
		PW cue	0.53 (0.07)**	0.51 (0.06)	<b>0.52 (0.07)</b>
		<b>M</b>	<b>0.57 (0.10)</b>	<b>0.51 (0.07)</b>	<i>0.54 (0.10)</i>
	Familiarised	W cue	0.61 (0.11)***	0.53 (0.09)*	<b>0.57 (0.11)</b>
		PW cue	0.53 (0.07)**	0.51 (0.10)	<b>0.52 (0.09)</b>
		<b>M</b>	<b>0.57 (0.10)</b>	<b>0.52 (0.10)</b>	<i>0.54 (0.10)</i>

Standard deviations are provided in parentheses. %CR = proportion of correct recognition; PW = pseudoword; W = word. Stars indicate significant differences from 0 for bias and gamma correlations and from 0.5 for AUROC2 (\*\* : p < .01; \*\*\* : p < .001).

## Annexe H. Chapitre 6 : Tableaux récapitulatif des moyennes de l'expérience 6.3.

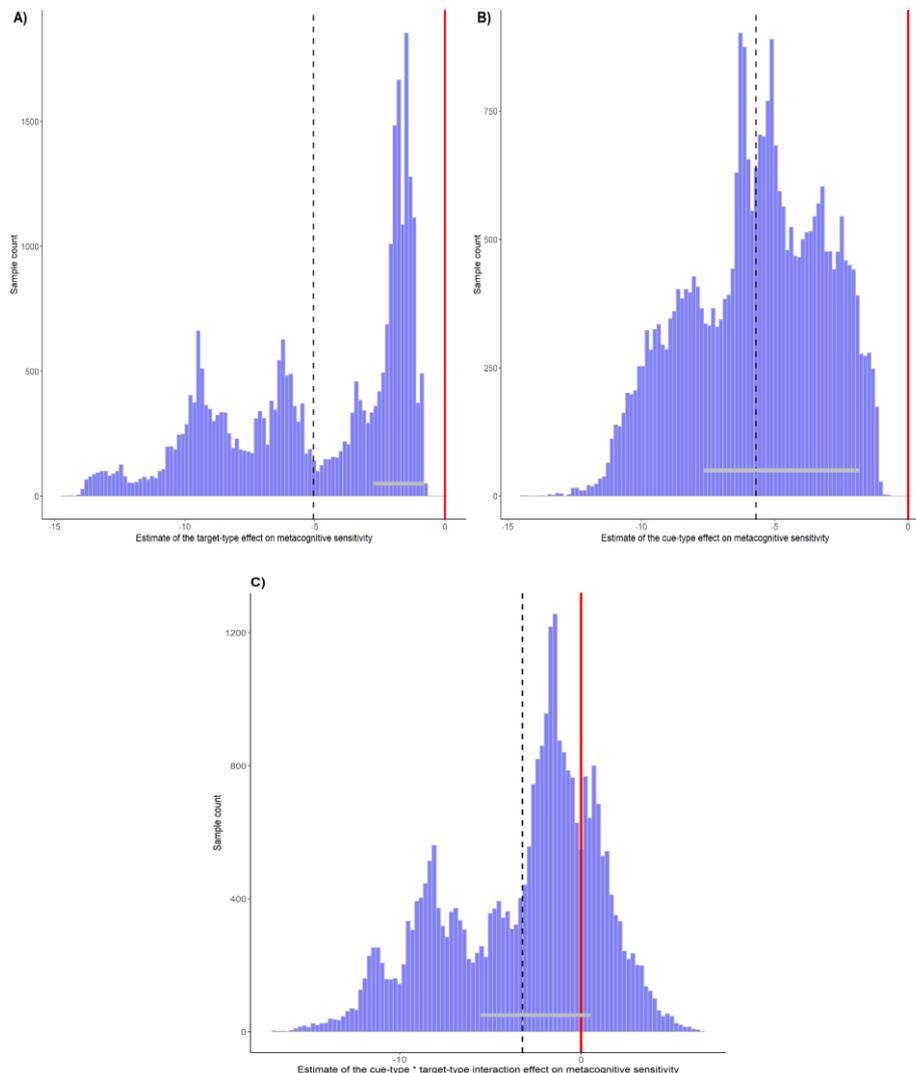
			W targets	PW targets	M
<b>Memory performance</b>					
Recall (%)		W cue	0.17 (0.16)	0.004 (0.01)	<b>0.09 (0.14)</b>
		PW cue	0.009 (0.02)	0 (0)	<b>0.005 (0.02)</b>
		<b>M</b>	<b>0.09 (0.14)</b>	<b>0.002 (0.01)</b>	
%CR	All trials	W cue	0.77 (0.12)	0.72 (0.09)	<b>0.74 (0.11)</b>
		PW cue	0.71 (0.10)	0.67 (0.10)	<b>0.69 (0.11)</b>
		<b>M</b>	<b>0.74 (0.12)</b>	<b>0.69 (0.10)</b>	
	Unrecalled	W cue	0.73 (0.12)	0.71 (0.09)	<b>0.72 (0.10)</b>
		PW cue	0.71 (0.10)	0.67 (0.10)	<b>0.69 (0.11)</b>
		<b>M</b>	<b>0.72 (0.11)</b>	<b>0.69 (0.10)</b>	
d' (recognition)	All trials	W cue	1.55 (0.75)	1.14 (0.55)	<b>1.34 (0.69)</b>
		PW cue	1.14 (0.66)	0.90 (0.57)	<b>1.02 (0.62)</b>
		<b>M</b>	<b>1.35 (0.73)</b>	<b>1.02 (0.57)</b>	
	Unrecalled	W cue	1.27 (0.67)	1.13 (0.54)	<b>1.20 (0.61)</b>
		PW cue	1.13 (0.65)	0.90 (0.57)	<b>1.02 (0.62)</b>
		<b>M</b>	<b>1.20 (0.66)</b>	<b>1.02 (0.56)</b>	
<b>Metacognitive performance</b>					
Magnitude (transformed values)	All trials	W cue	0.80 (0.07)	0.75 (0.07)	<b>0.78 (0.08)</b>
		PW cue	0.64 (0.08)	0.64 (0.08)	<b>0.64 (0.08)</b>
		<b>M</b>	<b>0.72 (0.11)</b>	<b>0.69 (0.10)</b>	
	Unrecalled	W cue	0.77 (0.07)	0.75 (0.07)	<b>0.76 (0.07)</b>
		PW cue	0.64 (0.08)	0.64 (0.08)	<b>0.64 (0.08)</b>
		<b>M</b>	<b>0.70 (0.10)</b>	<b>0.69 (0.10)</b>	
Gamma correlation	All trials	W cue	0.43 (0.34)***	0.20 (0.37)**	<b>0.31 (0.37)</b>
		PW cue	0.29 (0.47)**	0.27 (0.48)**	<b>0.28 (0.47)</b>
		<b>M</b>	<b>0.36 (0.41)</b>	<b>0.23 (0.43)</b>	
	Unrecalled	W cue	0.26 (0.41)**	0.20 (0.37)**	<b>0.23 (0.39)</b>
		PW cue	0.28 (0.47)**	0.27 (0.48)**	<b>0.27 (0.47)</b>
		<b>M</b>	<b>0.27 (0.44)</b>	<b>0.23 (0.43)</b>	
AUROC2	All trials	W cue	0.63 (0.11)***	0.55 (0.10)**	<b>0.59 (0.11)</b>
		PW cue	0.52 (0.09)	0.53 (0.09)	<b>0.52 (0.09)</b>
		<b>M</b>	<b>0.58 (0.11)</b>	<b>0.54 (0.09)</b>	
	Unrecalled	W cue	0.57 (0.11)**	0.55 (0.09)**	<b>0.56 (0.10)</b>
		PW cue	0.52 (0.09)	0.53 (0.09)	<b>0.52 (0.09)</b>
		<b>M</b>	<b>0.54 (0.10)</b>	<b>0.54 (0.09)</b>	

Standard deviations are provided in parentheses. %CR = proportion of correct recognition; PW = pseudoword; W = word. Stars indicate significant differences from 0 for bias and gamma correlations and from 0.5 for AUROC2 (\*\* :  $p < .01$ ; \*\*\* :  $p < .001$ ).

## Annexe I. Chapitre 6 : Résultats des analyses sur l'efficacité métacognitive.

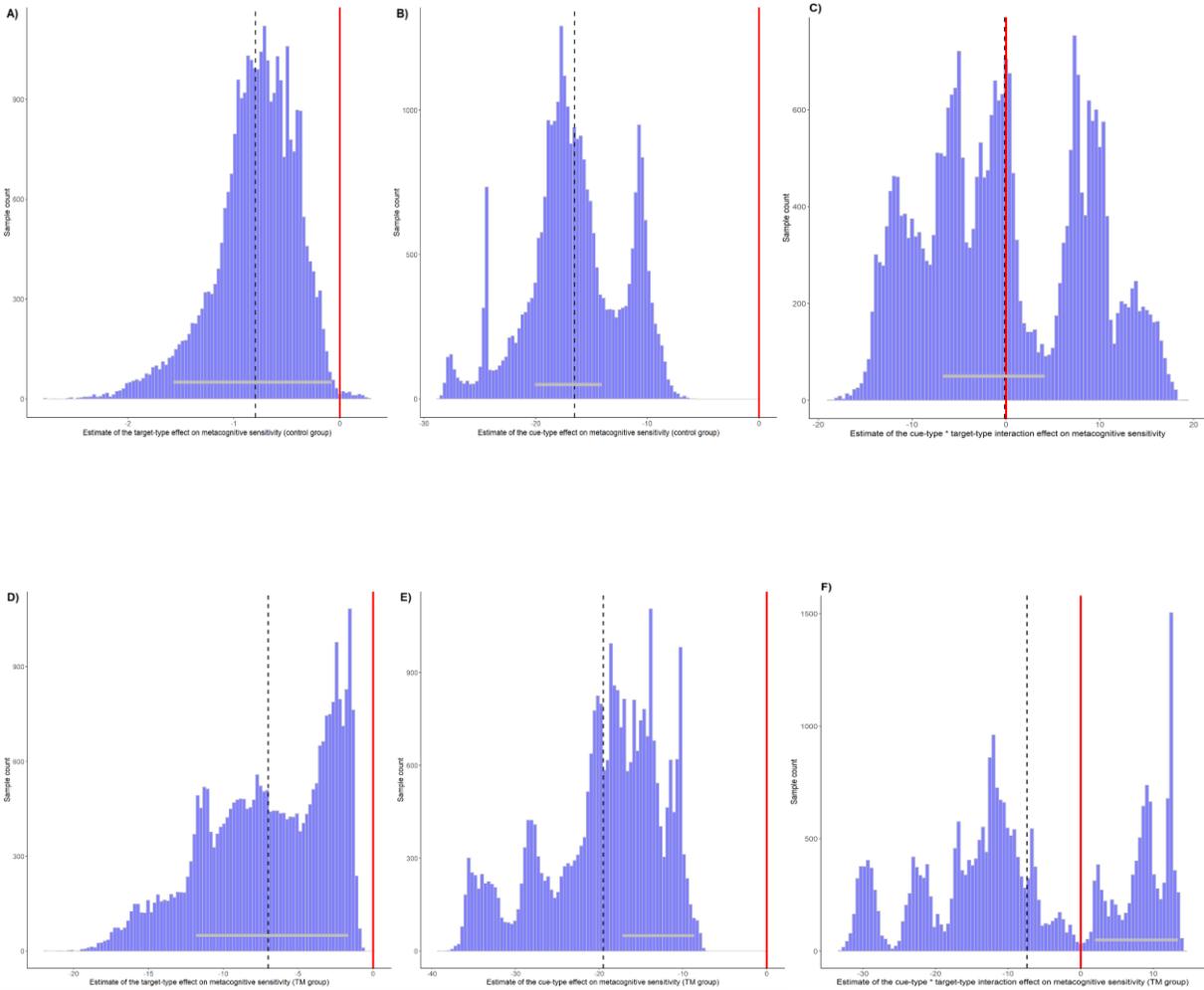
The hierarchical bayesian framework for estimating meta-d' was used as it has better reliability than other estimations procedures for designs with few trials (i.e., inferior to 100; Fleming, 2017). However, the three MCMC chains did not converge in none of the experiments, which could explain why the data does not follow a gaussian distribution. Although results show that HDI of estimates for the target-type and the cue-type effects do not overlap with zero (except in Experiment 4), which suggest significant effects, such interpretation is very limited because chains did not converge and data do not follow a normal distribution. Moreover, it also prevents comparison of HDI to determine the effect of groups in Experiment 2 and 3.

### Experiment 1.



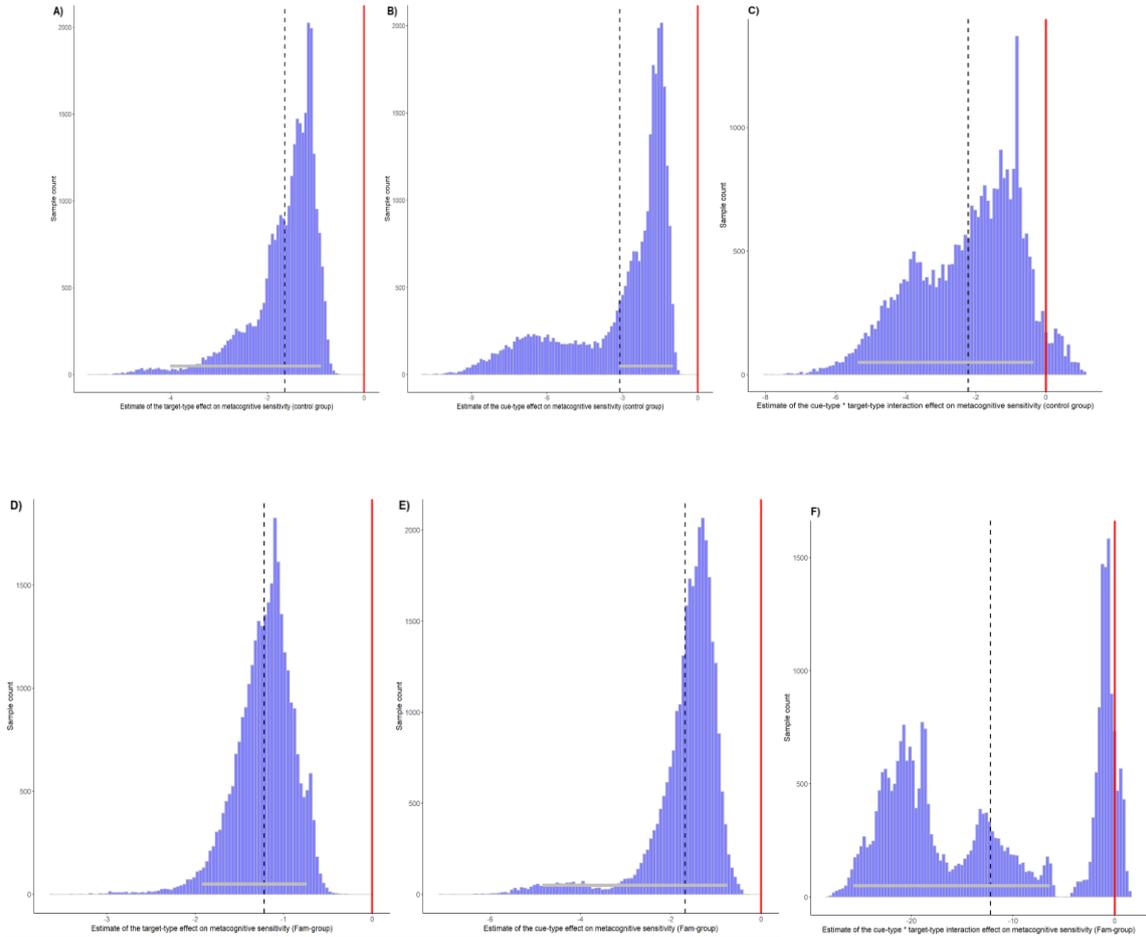
**Figure I.** Posterior distributions of mean ( $\mu$ , dashed line) effects of target-type (A), cue-type (B) and their interaction (C) in Experiment 6.1. The white lines are the 95% high density intervals (HDI). Overlaps with zero (red line) suggest no significant effect.

# Experiment 2.



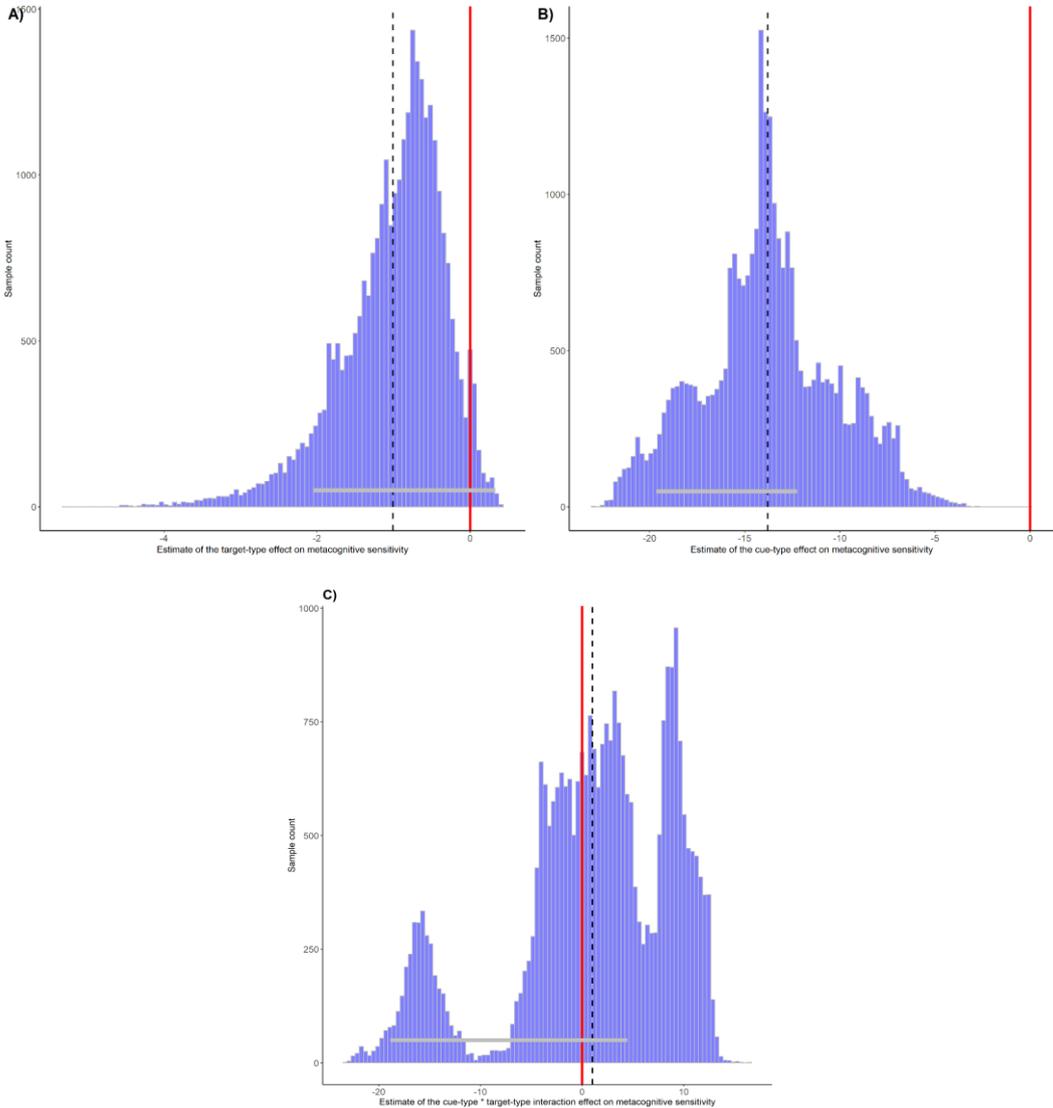
**Figure II.** Posterior distributions of mean ( $\mu$ , dashed line) effects of target-type, cue-type and their interaction in Experiment 6.2. Results from the control group are at the top (A, B, C) and results from the TM group are at the bottom (E, F, G). The white lines are the 95% high density intervals (HDI). Overlaps with zero (red line) suggest no significant effect.

# Experiment 3.

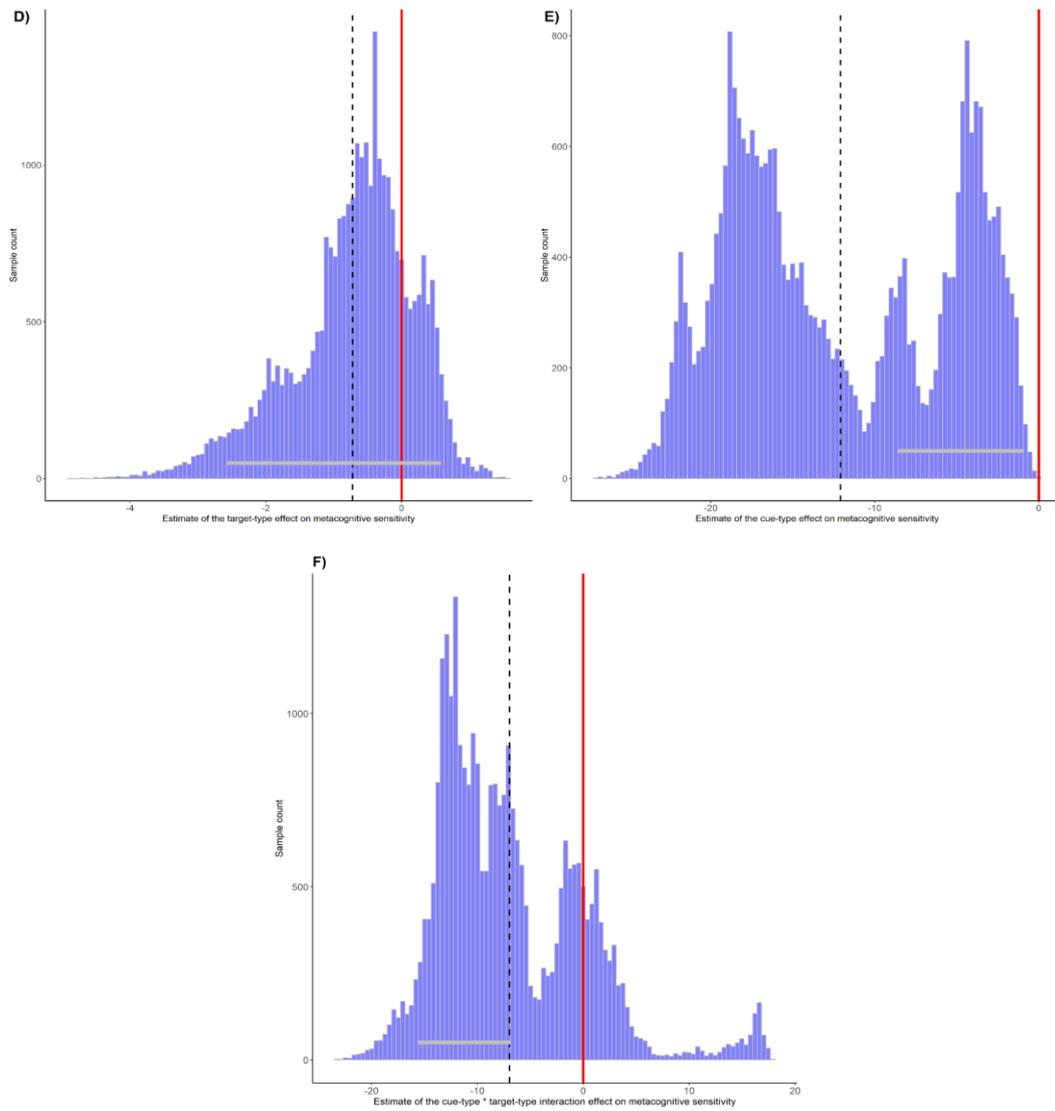


**Figure III.** Posterior distributions of mean ( $\mu$ , dashed line) effects of target-type, cue-type and their interaction in Experiment 6.3. Results from the control group are at the top (A, B, C) and results from the familiarised group are at the bottom (E, F, G). The white lines are the 95% high density intervals (HDI). Overlaps with zero (red line) suggest no significant effect.

# Experiment 4.



**Figure IV.** Posterior distributions of mean ( $\mu$ , dashed line) effects of target-type (A), cue-type (B) and their interaction (C) in Experiment 6.4 when all trials were included. The white lines are the 95% high density intervals (HDI). Overlaps with zero (red line) suggest no significant effect.



**Figure V.** Posterior distributions of mean ( $\mu$ , dashed line) effects of target-type (D), cue-type (E) and their interaction (F) in Experiment 4 when only considering trials with no recall. The white lines are the 95% high density intervals (HDI). Overlaps with zero (red line) suggest no significant effect.

## Annexe J. Chapitre 7 : Tableaux récapitulatif des résultats de l'expérience 7.1.

		Cible congruente	Cible incongruente	M
<b>Performance de premier ordre</b>				
%RC	Distracteur congruent	0.76 (0.09)	0.76 (0.09)	<b>0.76 (0.09)</b>
	Distracteur incongruent	0.78 (0.10)	0.76 (0.11)	<b>0.77 (0.10)</b>
	<b>M</b>	<b>0.77 (0.10)</b>	<b>0.77 (0.10)</b>	
d'	Distracteur congruent	1.47 (0.60)	1.48 (0.62)	<b>1.48 (0.61)</b>
	Distracteur incongruent	1.56 (0.67)	1.50 (0.69)	<b>1.53 (0.68)</b>
	<b>M</b>	<b>1.51 (0.64)</b>	<b>1.49 (0.65)</b>	
<b>Performance de second ordre</b>				
Magnitude	Distracteur congruent	0.77 (0.08)	0.77 (0.08)	<b>0.77 (0.08)</b>
	Distracteur incongruent	0.77 (0.08)	0.77 (0.08)	<b>0.77 (0.08)</b>
	<b>M</b>	<b>0.77 (0.08)</b>	<b>0.77 (0.08)</b>	
Biais	Distracteur congruent	0.02 (0.08)	0.006 (0.08)	<b>0.01 (0.08)</b>
	Distracteur incongruent	-0.009 (0.08)	0.003 (0.09)	<b>-0.003 (0.09)</b>
	<b>M</b>	<b>0.003 (0.08)</b>	<b>0.005 (0.08)</b>	
Corrélations Gamma	Distracteur congruent	0.63 (0.19)***	0.66 (0.22)***	<b>0.65 (0.21)</b>
	Distracteur incongruent	0.62 (0.26)***	0.62 (0.24)***	<b>0.62 (0.25)</b>
	<b>M</b>	<b>0.62 (0.23)</b>	<b>0.64 (0.23)</b>	
AUROC2	Distracteur congruent	0.71 (0.07)***	0.72 (0.09)***	<b>0.72 (0.08)</b>
	Distracteur incongruent	0.70 (0.09)***	0.70 (0.09)***	<b>0.70 (0.09)</b>
	<b>M</b>	<b>0.71 (0.08)</b>	<b>0.71 (0.09)</b>	

Les écart-types sont indiqués entre parenthèses. %RC = proportion de reconnaissance correcte;

Les astérisques indiquent des différences significatives de 0 pour les scores gamma, et de 0.5 pour les scores AUROC2 (i.e., jugements équivalents au hasard) ; \*\* :  $p < .01$ ; \*\*\* :  $p < .001$

# Bibliographie

---

- Ackerman, R., & Zalmanov, H. (2012). The persistence of the fluency–confidence association in problem solving. *Psychonomic Bulletin & Review*, 19(6), 1187-1192. <https://doi.org/10.3758/s13423-012-0305-z>
- Adlam, A.-L. R., Patterson, K., & Hodges, J. R. (2009). “I remember it as if it were yesterday” : Memory for recent events in patients with semantic dementia. *Neuropsychologia*, 47(5), 1344-1351. <https://doi.org/10.1016/j.neuropsychologia.2009.01.029>
- Agnew, S. K., & Morris, R. G. (1998). The heterogeneity of anosognosia for memory impairment in Alzheimer’s disease : A review of the literature and a proposed model. *Aging & Mental Health*, 2(1), 7-19.
- Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition*, 146, 377-386. <https://doi.org/10.1016/j.cognition.2015.10.006>
- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93(2), 203.
- Allen-Burge, R., & Storandt, M. (2000). Age equivalence in feeling-of-knowing experiences. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55(4), 214-223. <https://doi.org/10.1093/geronb/55.4.P214>
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13(3), 219-235. <https://doi.org/10.1177/1088868309341564>
- Alzheimer, A. (1907). Über eine eigenartige erkrankung der hirnrinde. *Zentralbl. Nervenhe. Psych.*, 18, 177-179.
- Anderson, B. A., & Peterson, S. A. (2022). Concreteness and levels of processing : A test of the dual-coding hypothesis using dynamic visual noise. *Memory*, 30(10), 1405-1420. <https://doi.org/10.1080/09658211.2022.2120622>
- Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, 81(1), 126-131. <https://doi.org/10.1037/h0027455>
- Atienza, M., Crespo-Garcia, M., & Cantero, J. L. (2011). Semantic congruence enhances memory of episodic associations : Role of theta oscillations. *Journal of Cognitive Neuroscience*, 23(1), 75-90. <https://doi.org/10.1162/jocn.2009.21358>
- Babinski, J. (1914). Contribution à l'étude des troubles mentaux dans l'hémiplégie organique cérébrale (anosognosie)[A contribution to the study of mental disturbances in organic cerebral hemiplegia (anosognosia)]. *Revue Neurologique (Pans)*, 1, 845-848.
- Bäckman, L., & Karlsson, T. (1985). The relation between level of general knowledge and feeling-of-knowing : An adult age study. *Scandinavian Journal of Psychology*, 26(1), 249-258. <https://doi.org/10.1111/j.1467-9450.1985.tb01162.x>
- Bäckman, L., & Lipinska, B. (1993). Monitoring of general knowledge : Evidence for preservation in early Alzheimer’s disease. *Neuropsychologia*, 31(4), 335-345. [https://doi.org/10.1016/0028-3932\(93\)90157-U](https://doi.org/10.1016/0028-3932(93)90157-U)
- Bacon, E., Danion, J.-M., Kauffmann-Muller, F., & Bruant, A. (2001). Consciousness in schizophrenia : A metacognitive approach to semantic memory. *Consciousness and Cognition*, 10(4), 473-484. <https://doi.org/10.1006/ccog.2001.0519>

- Bacon, E., Danion, J.-M., Kauffmann-Muller, F., Schelstraete, M.-A., Bruant, A., Sellal, F., & Grange, D. (1998). Confidence level and feeling of knowing for episodic and semantic memory : An investigation of lorazepam effects on metamemory. *Psychopharmacology*, 138(3-4), 318-325. <https://doi.org/10.1007/s002130050677>
- Bacon, E., & Izaute, M. (2009). Metacognition in schizophrenia : Processes underlying patients' reflections on their own episodic memory. *Biological Psychiatry*, 66(11), 1031-1037. <https://doi.org/10.1016/j.biopsych.2009.07.013>
- Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1413), 1345-1350. <https://doi.org/10.1098/rstb.2001.0957>
- Baker, L. (1991). Metacognition, reading, and science education. *Science learning: Processes and applications*, 2-13.
- Banks, S., & Weintraub, S. (2008). Self-awareness and self-monitoring of cognitive and behavioral deficits in behavioral variant frontotemporal dementia, primary progressive aphasia and probable Alzheimer's disease. *Brain and Cognition*, 67(1), 58-68. <https://doi.org/10.1016/j.bandc.2007.11.004>
- Baranski, J. V., & Petrusic, W. M. (1995). On the calibration of knowledge and perception. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 49(3), 397-407. <https://doi.org/10.1037/1196-1961.49.3.397>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577-660. <https://doi.org/10.1017/S0140525X99002149>
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.
- Bastin, C., Feyers, D., Souchay, C., Guillaume, B., Pepin, J.-L., Lemaire, C., Degueldre, C., Collette, F., & Salmon, E. (2012). Frontal and posterior cingulate metabolic impairment in the behavioral variant of frontotemporal dementia with impaired auto-noetic consciousness. *Human Brain Mapping*, 33(6), 1268-1278. <https://doi.org/10.1002/hbm.21282>
- Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, 28(5), 610-632. [https://doi.org/10.1016/0749-596X\(89\)90016-8](https://doi.org/10.1016/0749-596X(89)90016-8)
- Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the effect of semantic congruency on episodic memory : The role of integration and relatedness. *PLOS ONE*, 10(2), e0115624. <https://doi.org/10.1371/journal.pone.0115624>
- Belliard, S., Bon, L., LeMoal, S., Jonin, P. Y., Vercelletto, M., & LeBail, B. (2007). Semantic dementia. *Psychologie & NeuroPsychiatrie du vieillissement*, 5(2), 127-138. <https://doi.org/10.1684/pnv.2007.0080>
- Belliard, S., Jonin, P. Y., & Merck, C. (2010). Actualités sur la démence sémantique. *Revue de neuropsychologie*, 2(1), 31. <https://doi.org/10.3917/rne.021.0031>
- Belliard, S., Merck, C., Jonin, P. Y., Le Moal, S., & Vercelletto, M. (2011). Démence sémantique : Démographie et données neuropsychologiques initiales chez 82 patients. *Revue de neuropsychologie*, 3(4), 257. <https://doi.org/10.3917/rne.034.0257>
- Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, 31(2), 297-305. <https://doi.org/10.3758/BF03194388>

- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory : When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*(1), 55-68. <https://doi.org/10.1037/0096-3445.127.1.55>
- Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2022). Convicting with confidence? Why we should not over-rely on eyewitness confidence. *Memory*, *30*(1), 10-15. <https://doi.org/10.1080/09658211.2020.1849308>
- Bertrand, J. M., Mazancieux, A., Moulin, C. J. A., Béjot, Y., Rouaud, O., & Souchay, C. (2019). In the here and now : Short term memory predictions are preserved in Alzheimer's disease. *Cortex*, *119*, 158-164. <https://doi.org/10.1016/j.cortex.2019.03.027>
- Besken, M. (2016). Picture-perfect is not perfect for metamemory : Testing the perceptual fluency hypothesis with degraded images. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(9), 1417-1433. <https://doi.org/10.1037/xlm0000246>
- Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, *41*(6), 897-903. <https://doi.org/10.3758/s13421-013-0307-8>
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527-536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning : Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(1), 417-444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Bonitz, V. S., & Gordon, R. D. (2008). Attention to smoking-related and incongruous objects during scene viewing. *Acta Psychologica*, *129*(2), 255-263. <https://doi.org/10.1016/j.actpsy.2008.08.006>
- Bonner, M. F., Ash, S., & Grossman, M. (2010). The new classification of primary progressive aphasia into semantic, logopenic, or nonfluent/agrammatic variants. *Current Neurology and Neuroscience Reports*, *10*(6), 484-490. <https://doi.org/10.1007/s11910-010-0140-4>
- Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports*, *4*(1), 39-44. <https://doi.org/10.2466/pr0.1958.4.g.39>
- Boutoleau-Bretonnière, C., Evrard, C., Hardouin, J. B., Rocher, L., Charriau, T., Etcharry-Bouyx, F., Auriacombe, S., Richard-Mornas, A., Lebert, F., Pasquier, F., Sauvaget, A., Bulteau, S., Vercelletto, M., Derkinderen, P., Bretonnière, C., & Thomas-Antérion, C. (2015). DAPHNE : A new tool for the assessment of the behavioral variant of frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders Extra*, *5*(3), 503-516. <https://doi.org/10.1159/000440859>
- Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 529-533. [https://doi.org/10.1016/S0022-5371\(70\)80096-2](https://doi.org/10.1016/S0022-5371(70)80096-2)
- Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, *20*(2), 119-120. <https://doi.org/10.3758/BF03335632>
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*(9), 1207-1215. [https://doi.org/10.1016/S0028-3932\(00\)00034-8](https://doi.org/10.1016/S0028-3932(00)00034-8)

- Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in “associative” false memory. *Psychonomic Bulletin & Review*, *15*(6), 1035-1053. <https://doi.org/10.3758/PBR.15.6.1035>
- Brandt, M., Carvalho, R. L. S. de, Belfort, T., & Dourado, M. C. N. (2018). Metamemory monitoring in Alzheimer’s disease : A systematic review. *Dementia & Neuropsychologia*, *12*(4), 337-352. <https://doi.org/10.1590/1980-57642018dn12-040002>
- Brewer, G. A., Marsh, R. L., Clark-Foos, A., & Meeks, J. T. (2010). Noncriterial recollection influences metacognitive monitoring and control processes. *Quarterly Journal of Experimental Psychology*, *63*(10), 1936-1942. <https://doi.org/10.1080/17470210903551638>
- Brewer, W. F., & Sampaio, C. (2012). The metamemory approach to confidence : A test using semantic memory. *Journal of Memory and Language*, *67*(1), 59-77. <https://doi.org/10.1016/j.jml.2012.04.002>
- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. *Metacognition, motivation, and understanding*.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*(2), 204-223. <https://doi.org/10.1037/0033-2909.109.2.204>
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 325-337. [https://doi.org/10.1016/S0022-5371\(66\)80040-3](https://doi.org/10.1016/S0022-5371(66)80040-3)
- Bruce, P. R., Coyne, A. C., & Botwinick, J. (1982). Adult age differences in metamemory. *Journal of Gerontology*, *37*(3), 354-357. <https://doi.org/10.1093/geronj/37.3.354>
- Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, *24*(5), 654-663. <https://doi.org/10.1037/0012-1649.24.5.654>
- Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.
- Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*(1), 51-64. <https://doi.org/10.1037/xge0000505>
- Carroll, M., & Nelson, T. O. (1993). Effect of overlearning on the feeling of knowing is more detectable in within-subject than in between-subject designs. *The American Journal of Psychology*, *106*(2), 227. <https://doi.org/10.2307/1423169>
- Carroll, M., Nelson, T. O., & Kirwan, A. (1997). Tradeoff of semantic relatedness and degree of overlearning : Differential effects on metamemory and on long-term retention. *Acta Psychologica*, *95*(3), 239-253. [https://doi.org/10.1016/S0001-6918\(96\)00040-6](https://doi.org/10.1016/S0001-6918(96)00040-6)
- Castel, A. D., McCabe, D. P., & Roediger, H. L. (2007). Illusions of competence and overestimation of associative memory for identical items : Evidence from judgments of learning. *Psychonomic Bulletin & Review*, *14*(1), 107-111. <https://doi.org/10.3758/BF03194036>
- Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., Leschziner, G., Rossor, A. M., Stevens, J. M., Ciolotti, L., & Rossor, M. N. (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer’s disease. *Annals of Neurology*, *49*(4), 433-442. <https://doi.org/10.1002/ana.92>

- Chapman, C. A., Hasan, O., Schulz, P. E., & Martin, R. C. (2020). Evaluating the distinction between semantic knowledge and semantic access : Evidence from semantic dementia and comprehension-impaired stroke aphasia. *Psychonomic Bulletin & Review*, 27(4), 607-639. <https://doi.org/10.3758/s13423-019-01706-6>
- Chen, L., Lambon Ralph, M. A., & Rogers, T. T. (2017). A unified model of human semantic knowledge and its disorders. *Nature Human Behaviour*, 1(3), 0039. <https://doi.org/10.1038/s41562-016-0039>
- Chen, L., & Rogers, T. T. (2014). Revisiting domain-general accounts of category specificity in mind and brain : Revisiting domain-general accounts of category specificity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(3), 327-344. <https://doi.org/10.1002/wcs.1283>
- Chen, S., & McDunn, B. A. (2022). Metacognition : History, measurements, and the role in early childhood development and education. *Learning and Motivation*, 78, 101786. <https://doi.org/10.1016/j.lmot.2022.101786>
- Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses : A structural equation modeling approach. *Psychological Methods*, 19(2), 211-229. <https://doi.org/10.1037/a0032968>
- Chua, E. F., Pergolizzi, D., & Weintraub, R. R. (2014). The cognitive neuroscience of metamemory monitoring : Understanding metamemory processes, subjective levels expressed, and metacognitive accuracy. In S. M. Fleming & C. D. Frith (Éds.), *The Cognitive Neuroscience of Metacognition* (p. 267-291). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-45190-4\\_12](https://doi.org/10.1007/978-3-642-45190-4_12)
- Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural correlates of metamemory : A comparison of feeling-of-knowing and retrospective confidence judgments. *Journal of Cognitive Neuroscience*, 21(9), 1751-1765. <https://doi.org/10.1162/jocn.2009.21123>
- Clarke, F. R., Birdsall, T. G., & Tanner, W. P. (1959). Two types of ROC curves and definitions of parameters. *The Journal of the Acoustical Society of America*, 31(5), 629-630. <https://doi.org/10.1121/1.1907764>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240-247. [https://doi.org/10.1016/S0022-5371\(69\)80069-1](https://doi.org/10.1016/S0022-5371(69)80069-1)
- Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, 12(1), 50-71. <https://doi.org/10.1037/0882-7974.12.1.50>
- Corbett, F., Jefferies, E., Ehsan, S., & Lambon Ralph, M. A. (2009). Different impairments of semantic cognition in semantic dementia and semantic aphasia : Evidence from the non-verbal domain. *Brain*, 132(9), 2593-2608. <https://doi.org/10.1093/brain/awp146>
- Cosentino, S., Brickman, A. M., Griffith, E., Habeck, C., Cines, S., Farrell, M., Shaked, D., Huey, E. D., Briner, T., & Stern, Y. (2015). The right insula contributes to memory awareness in cognitively diverse older adults. *Neuropsychologia*, 75, 163-169. <https://doi.org/10.1016/j.neuropsychologia.2015.05.032>
- Cosentino, S., Metcalfe, J., Cary, M. S., De Leon, J., & Karlawish, J. (2011). Memory awareness influences everyday decision making capacity about medication management

- in Alzheimer's disease. *International Journal of Alzheimer's Disease*, 2011, 483897. <https://doi.org/10.4061/2011/483897>
- Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(1), 142-150. <https://doi.org/10.1037/0278-7393.18.1.142>
- Coyle-Gilchrist, I. T. S., Dick, K. M., Patterson, K., Vázquez Rodríguez, P., Wehmann, E., Wilcox, A., Lansdall, C. J., Dawson, K. E., Wiggins, J., Mead, S., Brayne, C., & Rowe, J. B. (2016). Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*, 86(18), 1736-1743. <https://doi.org/10.1212/WNL.0000000000002638>
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671-684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Crutch, S. J., & Warrington, E. K. (2002). Preserved calculation skills in a case of semantic dementia. *Cortex*, 38(3), 389-399. [https://doi.org/10.1016/S0010-9452\(08\)70667-1](https://doi.org/10.1016/S0010-9452(08)70667-1)
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17-22. <https://doi.org/10.1037/h0046671>
- Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence. *Law and Human Behavior*, 4(4), 243-260. <https://doi.org/10.1007/BF01040617>
- DeLozier, S. J., & Davalos, D. (2016). A systematic review of metacognitive differences between Alzheimer's disease and frontotemporal dementia. *American Journal of Alzheimer's Disease & Other Dementias*, 31(5), 381-388. <https://doi.org/10.1177/1533317515618899>
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231-264. <https://doi.org/10.1007/s11409-008-9029-x>
- Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, 22(1), 122-133. <https://doi.org/10.1037/0882-7974.22.1.122>
- Douchemane, D., Isingrini, M., & Souchay, C. (2007). Vieillesse, fonctions exécutives et mémoire épisodique et en mémoire sémantique. *L'Année Psychologique*, 107(04), 597. <https://doi.org/10.4074/S0003503307004046>
- Dougherty, M. R., Robey, A. M., & Buttaccio, D. (2018). Do metacognitive judgments alter memory performance beyond the benefits of retrieval practice? A comment on and replication attempt of Dougherty, Scheck, Nelson, and Narens (2005). *Memory & Cognition*, 46(4), 558-565. <https://doi.org/10.3758/s13421-018-0791-y>
- Dougherty, M. R., Scheck, P., Nelson, T. O., & Narens, L. (2005). Using the past to predict the future. *Memory & Cognition*, 33(6), 1096-1115. <https://doi.org/10.3758/BF03193216>
- Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: A Frontal Assessment Battery at bedside. *Neurology*, 55(11), 1621-1626.

- Dunlosky, J., & Bjork, R. A. (2014). The integrated nature of metamemory and memory. In *Handbook of Metamemory and Memory*. Routledge. <https://doi.org/10.4324/9780203805503.ch2>
- Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging, 15*(3), 462-474. <https://doi.org/10.1037/0882-7974.15.3.462>
- Dunlosky, J., & Matvey, G. (2001). Empirical analysis of the intrinsic–extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1180-1191. <https://doi.org/10.1037/0278-7393.27.5.1180>
- Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*(4), 374-380. <https://doi-org.insb.bib.cnrs.fr/10.3758/BF03210921>
- Dunlosky, J., & Tauber, S. K. (2014). Understanding people's metacognitive judgments: An isomechanism framework and its implications for applied and theoretical research. In T. J. Perfect & D. S. Lindsay, *The SAGE Handbook of Applied Memory* (p. 444-464). SAGE Publications Ltd. <https://doi.org/10.4135/9781446294703.n25>
- Duval, C., Desgranges, B., de La Sayette, V., Belliard, S., Eustache, F., & Piolino, P. (2012). What happens to personal identity when semantic knowledge degrades? A study of the self and autobiographical memory in semantic dementia. *Neuropsychologia, 50*(2), 254-265. <https://doi.org/10.1016/j.neuropsychologia.2011.11.019>
- Eakin, D. K., & Hertzog, C. (2012). Age invariance in feeling of knowing during implicit interference effects. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 67*(5), 555-562. <https://doi.org/10.1093/geronb/gbr150>
- Eakin, D. K., Hertzog, C., & Harris, W. (2014). Age invariance in semantic and episodic metamemory: Both younger and older adults provide accurate feeling-of-knowing for names of faces. *Aging, Neuropsychology, and Cognition, 21*(1), 27-51. <https://doi.org/10.1080/13825585.2013.775217>
- Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*(1), 3-14. <https://doi.org/10.1016/j.edurev.2005.11.001>
- Elman, J. A., Klostermann, E. C., Marian, D. E., Verstaen, A., & Shimamura, A. P. (2012). Neural correlates of metacognitive monitoring during episodic and semantic retrieval. *Cognitive, Affective, & Behavioral Neuroscience, 12*(3), 599-609. <https://doi.org/10.3758/s13415-012-0096-8>
- Emanuel Robinson, A., Hertzog, C., & Dunlosky, J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition, 13*(3-4), 458-478. <https://doi.org/10.1080/13825580600572983>
- Epstein, M. L., Phillips, W. D., & Johnson, S. J. (1975). Recall of related and unrelated word pairs as a function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 1*(2), 149-152. <https://doi.org/10.1037/0278-7393.1.2.149>
- Ernst, A., Moulin, C. J. A., Souchay, C., Mograbi, D. C., & Morris, R. (2015). *Anosognosia and metacognition in Alzheimer's disease* (J. Dunlosky & S. (Uma) K. Tauber, Éd.s.; Vol. 1). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199336746.013.12>

- Eslinger, P. J. (2005). Metacognitive deficits in frontotemporal dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(12), 1630-1635. <https://doi.org/10.1136/jnnp.2004.053157>
- Evans, S., & Azzopardi, P. (2007). Evaluation of a « bias-free » measure of awareness. *Spatial Vision*, 20(1-2), 61-77. <https://doi.org/10.1163/156856807779369742>
- Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *The Journal of Neuroscience*, 38(2), 263-277. <https://doi.org/10.1523/JNEUROSCI.0322-17.2017>
- Farah, M. J., Hammond, K. M., Mehta, Z., & Ratcliff, G. (1989). Category-specificity and modality-specificity in semantic memory. *Neuropsychologia*, 27(2), 193-200. [https://doi.org/10.1016/0028-3932\(89\)90171-1](https://doi.org/10.1016/0028-3932(89)90171-1)
- Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2), 133-146. <https://doi.org/10.1111/1469-8986.3920133>
- Finn, B. (2008). Framing effects on metacognitive monitoring and control. *Memory & Cognition*, 36(4), 813-821. <https://doi.org/10.3758/MC.36.4.813>
- Fitzgerald, L. M., Arvaneh, M., & Dockree, P. M. (2017). Domain-specific and domain-general processes underlying metacognitive judgments. *Consciousness and Cognition*, 49, 264-277. <https://doi.org/10.1016/j.concog.2017.01.011>
- Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development*, 14(4), 272-278. <https://doi.org/10.1159/000271221>
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. *The nature of intelligence*, 231-235.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleming, S. M. (2017). HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, 2017(1), nix007. <https://doi.org/10.1093/nc/nix007>
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1338-1349. <https://doi.org/10.1098/rstb.2011.0417>
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8, 443. <https://doi.org/10.3389/fnhum.2014.00443>
- Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J.-C. (2016). Metacognition about the past and future: Quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, 2016(1), niw018. <https://doi.org/10.1093/nc/niw018>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". *Journal of Psychiatric Research*, 12(3), 189-198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599), 2213-2223. <https://doi.org/10.1098/rstb.2012.0123>

- Gainotti, G. (2017). The differential contributions of conceptual representation format and language structure to levels of semantic abstraction capacity. *Neuropsychology Review*, 27(2), 134-145. <https://doi.org/10.1007/s11065-016-9339-8>
- Gallo, D. A. (2006). *Associative memory illusions*. New York: Psychology Press.
- Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10(4), 843-876. <https://doi.org/10.3758/BF03196546>
- Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition*, 18(1), 23-30. <https://doi.org/10.3758/BF03202642>
- Geraci, L., McCabe, D. P., & Guillory, J. J. (2009). On interpreting the relationship between remember-know judgments and confidence: The role of instructions. *Consciousness and Cognition*, 18(3), 701-709. <https://doi.org/10.1016/j.concog.2009.04.010>
- Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition*, 66(2), 115-152. [https://doi.org/10.1016/S0010-0277\(98\)00020-1](https://doi.org/10.1016/S0010-0277(98)00020-1)
- Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, 49(268), 732. <https://doi.org/10.2307/2281536>
- Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W., & Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, 55(3), 335-346. <https://doi.org/10.1002/ana.10825>
- Gotts, S. J., & Plaut, D. C. (2002). The impact of synaptic depression following brain damage: A connectionist account of « access/refractory » and « degraded-store » semantic impairments. *Cognitive, Affective, & Behavioral Neuroscience*, 2(3), 187-213. <https://doi.org/10.3758/CABN.2.3.187>
- Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia*, 38(3), 313-324.
- Grainger, C., Williams, D. M., & Lind, S. E. (2014). Metacognition, metamemory, and mindreading in high-functioning adults with autism spectrum disorder. *Journal of Abnormal Psychology*, 123(3), 650-659. <https://doi.org/10.1037/a0036531>
- Grainger, C., Williams, D. M., & Lind, S. E. (2016). Judgment of learning accuracy in high-functioning adolescents and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46(11), 3570-3582. <https://doi.org/10.1007/s10803-016-2895-1>
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.
- Greene, M. R., Botros, A. P., Beck, D. M., & Fei-Fei, L. (2015). What you see is what you expect: Rapid scene understanding benefits from prior experience. *Attention, Perception, & Psychophysics*, 77(4), 1239-1251. <https://doi.org/10.3758/s13414-015-0859-8>
- Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. *Neuroscience & Biobehavioral Reviews*, 55, 88-97. <https://doi.org/10.1016/j.neubiorev.2015.04.006>

- Grossman, M. (2010). Primary progressive aphasia : Clinicopathological correlations. *Nature Reviews Neurology*, 6(2), 88-97. <https://doi.org/10.1038/nrneurol.2009.216>
- Guggenmos, M. (2021). Measuring metacognitive performance: Type 1 performance dependence and test-retest reliability. *Neuroscience of Consciousness*, 2021(1), niab040. <https://doi.org/10.1093/nc/niab040>
- Gutchess, A. H., & Park, D. C. (2009). Effects of ageing on associative memory for related and unrelated pictures. *European Journal of Cognitive Psychology*, 21(2-3), 235-254. <https://doi.org/10.1080/09541440802257274>
- Haaf, J. M., Rhodes, S., Naveh-Benjamin, M., Sun, T., Snyder, H. K., & Rouder, J. N. (2021). Revisiting the remember-know task : Replications of Gardiner and Java (1990). *Memory & Cognition*, 49(1), 46-66. <https://doi.org/10.3758/s13421-020-01073-x>
- Halamish, V., McGillivray, S., & Castel, A. D. (2011). Monitoring one's own forgetting in younger and older adults. *Psychology and Aging*, 26(3), 631-635. <https://doi.org/10.1037/a0022852>
- Hanczakowski, M., Zawadzka, K., & Cockcroft-McKay, C. (2014). Feeling of knowing and restudy choices. *Psychonomic Bulletin & Review*, 21(6), 1617-1622. <https://doi.org/10.3758/s13423-014-0619-0>
- Hannesdottir, K., & Morris, R. G. (2007). Primary and secondary anosognosia for memory impairment in patients with Alzheimer's disease. *Cortex*, 43(7), 1020-1030.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading : Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662-720. <https://doi.org/10.1037/0033-295X.111.3.662>
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56(4), 208-216. <https://doi.org/10.1037/h0022263>
- Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, 6(5), 685-691. [https://doi.org/10.1016/S0022-5371\(67\)80072-0](https://doi.org/10.1016/S0022-5371(67)80072-0)
- Heikkilä, J., Alho, K., Hyvönen, H., & Tiippana, K. (2015). Audiovisual semantic congruency during encoding enhances memory performance. *Experimental Psychology*, 62(2), 123-130. <https://doi.org/10.1027/1618-3169/a000279>
- Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(1), 22-34. <https://doi.org/10.1037/0278-7393.29.1.22>
- Hertzog, C., Dunlosky, J., & Sinclair, S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. *Memory & Cognition*, 38(6), 771-784. <https://doi.org/10.3758/MC.38.6.771>
- Hertzog, C., Fulton, E. K., Sinclair, S. M., & Dunlosky, J. (2014). Recalled aspects of original encoding strategies influence episodic feelings of knowing. *Memory & Cognition*, 42(1), 126-140. <https://doi.org/10.3758/s13421-013-0348-z>
- Hertzog, C., Kidder, D. P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning : Is monitoring accuracy spared or impaired? *Psychology and Aging*, 17(2), 209-225. <https://doi.org/10.1037/0882-7974.17.2.209>
- Hertzog, C., & Touron, D. R. (2011). Age differences in memory retrieval shift : Governed by feeling-of-knowing? *Psychology and Aging*, 26(3), 647-660. <https://doi.org/10.1037/a0021875>

- Heyes, C., Bang, D., Shea, N., Frith, C. D., & Fleming, S. M. (2020). Knowing ourselves together : The cultural origins of metacognition. *Trends in Cognitive Sciences*, 24(5), 349-362. <https://doi.org/10.1016/j.tics.2020.02.007>
- Hicks, J. L., & Marsh, R. L. (2002). On predicting the future states of awareness for recognition of unrecalable items. *Memory & Cognition*, 30(1), 60-66. <https://doi.org/10.3758/BF03195265>
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558. <https://doi.org/10.1002/sim.1186>
- Higham, P. A., Zawadzka, K., & Hanczakowski, M. (2016). Internal mapping and its impact on measures of absolute and relative metacognitive accuracy. *The Oxford handbook of metamemory*, 39-61.
- Hodges, J. R. (2000). The role of conceptual knowledge in object use : Evidence from semantic dementia. *Brain*, 123(9), 1913-1925. <https://doi.org/10.1093/brain/123.9.1913>
- Hodges, J. R., & Graham, K. S. (2001). Episodic memory : Insights from semantic dementia. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1413), 1423-1434. <https://doi.org/10.1098/rstb.2001.0943>
- Hodges, J. R., Mitchell, J., Dawson, K., Spillantini, M. G., Xuereb, J. H., McMonagle, P., Nestor, P. J., & Patterson, K. (2010). Semantic dementia : Demography, familial factors and survival in a consecutive series of 100 cases. *Brain*, 133(1), 300-306. <https://doi.org/10.1093/brain/awp248>
- Hodges, J. R., & Patterson, K. (2007). Semantic dementia : A unique clinicopathological syndrome. *The Lancet Neurology*, 6(11), 1004-1014. [https://doi.org/10.1016/S1474-4422\(07\)70266-1](https://doi.org/10.1016/S1474-4422(07)70266-1)
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia : Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115(6), 1783-1806. <https://doi.org/10.1093/brain/115.6.1783>
- Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, 7(1-3), 213-235. <https://doi.org/10.1080/135062800394775>
- Hosey, L. A., Peynircioğlu, Z. F., & Rabinovitz, B. E. (2009). Feeling of knowing for names in response to faces. *Acta Psychologica*, 130(3), 214-224. <https://doi.org/10.1016/j.actpsy.2008.12.007>
- Hou, C. E., Miller, B. L., & Kramer, J. H. (2005). Patterns of autobiographical memory loss in dementia. *International Journal of Geriatric Psychiatry*, 20(9), 809-815. <https://doi.org/10.1002/gps.1361>
- Hoven, M., Lebreton, M., Engelmann, J. B., Denys, D., Luigjes, J., & van Holst, R. J. (2019). Abnormalities of confidence in psychiatry : An overview and future perspectives. *Translational Psychiatry*, 9(1), 268. <https://doi.org/10.1038/s41398-019-0602-7>
- Howard, C. E., Andrés, P., Broks, P., Noad, R., Sadler, M., Coker, D., & Mazzoni, G. (2010). Memory, metamemory and their dissociation in temporal lobe epilepsy. *Neuropsychologia*, 48(4), 921-932. <https://doi.org/10.1016/j.neuropsychologia.2009.11.011>
- Hu, X., Li, T., Zheng, J., Su, N., Liu, Z., & Luo, L. (2015). How much do metamemory beliefs contribute to the font-size effect in judgments of learning? *PLOS ONE*, 10(11), e0142351. <https://doi.org/10.1371/journal.pone.0142351>

- Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, *193*, 104012. <https://doi.org/10.1016/j.cognition.2019.104012>
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words : Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685-701. [https://doi.org/10.1016/0749-596X\(91\)90032-F](https://doi.org/10.1016/0749-596X(91)90032-F)
- Irish, M. (2020). On the interaction between episodic and semantic representations – constructing a unified account of imagination. In A. Abraham (Éd.), *The Cambridge Handbook of the Imagination* (1<sup>re</sup> éd., p. 447-465). Cambridge University Press. <https://doi.org/10.1017/9781108580298.027>
- Irish, M., Bunk, S., Tu, S., Kamminga, J., Hodges, J. R., Hornberger, M., & Piguet, O. (2016). Preservation of episodic memory in semantic dementia : The importance of regions beyond the medial temporal lobes. *Neuropsychologia*, *81*, 50-60. <https://doi.org/10.1016/j.neuropsychologia.2015.12.005>
- Irish, M., Hornberger, M., Lah, S., Miller, L., Pengas, G., Nestor, P. J., Hodges, J. R., & Piguet, O. (2011). Profiles of recent autobiographical memory retrieval in semantic dementia, behavioural-variant frontotemporal dementia, and Alzheimer's disease. *Neuropsychologia*, *49*(9), 2694-2702. <https://doi.org/10.1016/j.neuropsychologia.2011.05.017>
- Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43-49. <https://doi.org/10.1016/j.cobeha.2020.01.016>
- Isingrini, M., Perrotin, A., & Souchay, C. (2008). Chapter 24 Aging, metamemory regulation and executive functioning. In *Progress in Brain Research* (Vol. 169, p. 377-392). Elsevier. [https://doi.org/10.1016/S0079-6123\(07\)00024-6](https://doi.org/10.1016/S0079-6123(07)00024-6)
- Isingrini, M., Sacher, M., Perrotin, A., Tacconnat, L., Souchay, C., Stoehr, H., & Bouazzaoui, B. (2016). Episodic feeling-of-knowing relies on noncriterial recollection and familiarity : Evidence using an online remember-know procedure. *Consciousness and Cognition*, *41*, 31-40. <https://doi.org/10.1016/j.concog.2016.01.011>
- Ivanoiu, A., Cooper, J. M., Shanks, M. F., & Venneri, A. (2006). Patterns of impairment in autobiographical memory in the degenerative dementias constrain models of memory. *Neuropsychologia*, *44*(10), 1936-1955. <https://doi.org/10.1016/j.neuropsychologia.2006.01.030>
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306-340. <https://doi.org/10.1037/0096-3445.110.3.306>
- Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory : Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology*, *17*(1), 3-11. <https://doi.org/10.3758/BF03337811>
- Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1). <https://doi.org/10.7771/1932-6246.1167>
- JASP team. (2020). *JASP* (Version 0.13.1) [Computer software].
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia : A case-series comparison. *Brain*, *129*(8), 2132-2147. <https://doi.org/10.1093/brain/awl153>

- Jeffreys, H. (1998). *The theory of probability*. Oxford University Press Oxford.
- Jemstedt, A., Schwartz, B. L., & Jönsson, F. U. (2018). Ease-of-learning judgments are based on both processing fluency and beliefs. *Memory*, 26(6), 807-815. <https://doi.org/10.1080/09658211.2017.1410849>
- Jönsson, F. U., Olsson, H., & Olsson, M. J. (2005). Odor emotionality affects the confidence in odor naming. *Chemical Senses*, 30(1), 29-35. <https://doi.org/10.1093/chemse/bjh254>
- Jönsson, F. U., & Olsson, M. J. (2003). Olfactory metacognition. *Chemical Senses*, 28(7), 651-658. <https://doi.org/10.1093/chemse/bjg058>
- Kalafat, M., Hugonot-Diener, L., & Poitrenaud, J. (2003). Standardisation et étalonnage français du « Mini Mental State » (MMS) version GRÉCO [French standardization and range for the GRECO version of the « Mini Mental State » (MMS)]. *Revue de Neuropsychologie*, 13(2), 209–236.
- Kan, I. P., Alexander, M. P., & Verfaellie, M. (2009). Contribution of prior semantic knowledge to new episodic learning in amnesia. *Journal of Cognitive Neuroscience*, 21(5), 938-944. <https://doi.org/10.1162/jocn.2009.21066>
- Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition : Evidence against a general metacognitive ability. *Memory & Cognition*, 28(1), 92-107. <https://doi.org/10.3758/BF03211579>
- Kelley, C. M. ., & Lindsay, D. S. (1993). Remembering mistaken for knowing : Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32(1), 1-24. <https://doi.org/10.1006/jmla.1993.1001>
- Kentridge, R. W., & Heywood, C. A. (2000). Metacognition and awareness. *Consciousness and Cognition*, 9(2), 308-312. <https://doi.org/10.1006/ccog.2000.0448>
- Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, 84(6), 1329-1342. <https://doi.org/10.1016/j.neuron.2014.12.015>
- Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, 324(5928), 759-764. <https://doi.org/10.1126/science.1169405>
- Kikyo, H., & Miyashita, Y. (2004). Temporal lobe activations of “feeling-of-knowing” induced by face-name associations. *NeuroImage*, 23(4), 1348-1357. <https://doi.org/10.1016/j.neuroimage.2004.08.013>
- Kikyo, H., Ohki, K., & Miyashita, Y. (2002). Neural correlates for feeling-of-knowing. *Neuron*, 36(1), 177-186. [https://doi.org/10.1016/S0896-6273\(02\)00939-X](https://doi.org/10.1016/S0896-6273(02)00939-X)
- Knopman, D. S., & Roberts, R. O. (2011). Estimating the number of persons with frontotemporal lobar degeneration in the US population. *Journal of Molecular Neuroscience*, 45(3), 330-335. <https://doi.org/10.1007/s12031-011-9538-y>
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, 100(4), 609-639. <https://doi.org/10.1037/0033-295X.100.4.609>
- Koriat, A. (1994). Memory's knowledge of its own knowledge : The accessibility account of the feeling of knowing. *Metacognition: Knowing about knowing*, 115-135.

- Koriat, A. (1995). Dissociating knowing and the feeling of knowing : Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, 124(3), 311-333. <https://doi.org/10.1037/0096-3445.124.3.311>
- Koriat, A. (1997). Monitoring one's own knowledge during study : A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349-370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A. (2000). The feeling of knowing : Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9(2), 149-171. <https://doi.org/10.1006/ccog.2000.0433>
- Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Éds.), *The Cambridge Handbook of Consciousness* (p. 289-326). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816789.012>
- Koriat, A. (2008). When confidence in a choice is independent of which choice is made. *Psychonomic Bulletin & Review*, 15(5), 997-1001. <https://doi.org/10.3758/PBR.15.5.997>
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting : The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, 133(4), 643-656. <https://doi.org/10.1037/0096-3445.133.4.643>
- Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments : Information-based and experience-based monitoring of one's own knowledge. In *Dual-process theories in social psychology* (p. 483-502). The Guilford Press.
- Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 34-53.
- Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1095-1105. <https://doi.org/10.1037/0278-7393.29.6.1095>
- Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments : Evidence from subjective confidence. *A handbook of memory and metamemory*, 117-136.
- Kralik, J. D., Lee, J. H., Rosenbloom, P. S., Jackson, P. C., Epstein, S. L., Romero, O. J., Sanz, R., Larue, O., Schmidtke, H. R., Lee, S. W., & McGreggor, K. (2018). Metacognition for a common model of cognition. *Procedia Computer Science*, 145, 730-739. <https://doi.org/10.1016/j.procs.2018.11.046>
- Krebs, R. M., Boehler, C. N., De Belder, M., & Egner, T. (2015). Neural conflict-control mechanisms improve memory for target stimuli. *Cerebral Cortex*, 25(3), 833-843. <https://doi.org/10.1093/cercor/bht283>
- Kruschke, J. K. (2015). *Doing Bayesian data analysis : A tutorial with R, JAGS, and Stan* (2<sup>e</sup> éd.). Academic Press.
- Kuhlmann, B. G., & Bayen, U. J. (2015). *Metacognitive Aspects of Source Monitoring* (J. Dunlosky & S. (Uma) K. Tauber, Éds.; Vol. 1). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199336746.013.8>
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463-470. [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6)

- Lachman, J. L., Lachman, R., & Thronesbery, C. (1979). Metamemory through the adult life span. *Developmental Psychology*, 15(5), 543-551. <https://doi.org/10.1037/0012-1649.15.5.543>
- Lai, E. R. (2011). Metacognition : A literature review. *Pearson Assessments Research Reports*, 41.
- Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120392. <https://doi.org/10.1098/rstb.2012.0392>
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42-55. <https://doi.org/10.1038/nrn.2016.150>
- Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory : Insights from semantic dementia. *Annals of the New York Academy of Sciences*, 1124(1), 61-76. <https://doi.org/10.1196/annals.1440.006>
- LaPointe, M. R. P., Lupianez, J., & Milliken, B. (2013). Context congruency effects in change detection: Opposing effects on detection and identification. *Visual Cognition*, 21(1), 99-122. <https://doi.org/10.1080/13506285.2013.787133>
- LaPointe, M. R. P., & Milliken, B. (2016). Semantically incongruent objects attract eye gaze when viewing scenes for change. *Visual Cognition*, 24(1), 63-77. <https://doi.org/10.1080/13506285.2016.1185070>
- Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, 15(8), 365-373. <https://doi.org/10.1016/j.tics.2011.05.009>
- Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 464-467.
- Levelt, W. J. M. (1992). Accessing words in speech production : Stages, processes and representations. *Cognition*, 42(1-3), 1-22. [https://doi.org/10.1016/0010-0277\(92\)90038-J](https://doi.org/10.1016/0010-0277(92)90038-J)
- Levelt, W. J. M. (2001). Spoken word production : A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98(23), 13464-13471. <https://doi.org/10.1073/pnas.231459498>
- Lipinska, B., & Bäckman, L. (1996). Feeling-of-knowing in fact retrieval : Further evidence for preservation in early Alzheimer's disease. *Journal of the International Neuropsychological Society*, 2(4), 350-358. <https://doi.org/10.1017/S1355617700001375>
- Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing : A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 44(1), 87-112. <https://doi.org/10.1037/h0084237>
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), 565-572. <https://doi.org/10.1037/0096-1523.4.4.565>
- Lupker, S. J., Harbluk, J. L., & Patrick, A. S. (1991). Memory for things forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(5), 897-907. <https://doi.org/10.1037/0278-7393.17.5.897>

- MacLaverly, S. N., & Hertzog, C. (2009). Do age-related differences in episodic feeling of knowing accuracy depend on the timing of the judgement? *Memory*, 17(8), 860-873. <https://doi.org/10.1080/09658210903374537>
- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.
- Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data : Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, 98(1), 185-199. <https://doi.org/10.1037/0033-2909.98.1.185>
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87(3), 252-271. <https://doi.org/10.1037/0033-295X.87.3.252>
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, 21(1), 422-430. <https://doi.org/10.1016/j.concog.2011.09.021>
- Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of Type 1 and Type 2 data : Meta-d', response-specific Meta-d', and the unequal variance SDT model. In S. M. Fleming & C. D. Frith (Éds.), *The Cognitive Neuroscience of Metacognition* (p. 25-66). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-45190-4\\_3](https://doi.org/10.1007/978-3-642-45190-4_3)
- Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L., & Schacter, D. L. (2003). Feeling-of-knowing in episodic memory : An event-related fMRI study. *NeuroImage*, 18(4), 827-836. [https://doi.org/10.1016/S1053-8119\(03\)00014-4](https://doi.org/10.1016/S1053-8119(03)00014-4)
- Marquié, J. C., & Huet, N. (2000). Age differences in feeling-of-knowing and confidence judgments as a function of knowledge domain. *Psychology and Aging*, 15(3), 451-461. <https://doi.org/10.1037/0882-7974.15.3.451>
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58(1), 25-45. <https://doi.org/10.1146/annurev.psych.57.102904.190143>
- Martin, A. (2016). GRAPES—Grounding representations in action, perception, and emotion systems : How object properties and categories are represented in the human brain. *Psychonomic Bulletin & Review*, 23(4), 979-990. <https://doi.org/10.3758/s13423-015-0842-3>
- Martinez, M. E. (2006). What is metacognition? *Phi Delta Kappan*, 87(9), 696-699. <https://doi.org/10.1177/003172170608700916>
- Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association : Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(2), 509-527. <https://doi.org/10.1037/a0014876>
- Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs) : An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29(2), 222-233. <https://doi.org/10.3758/BF03194916>
- Matvey, G., Dunlosky, J., & Schwartz, B. (2006). The effects of categorical relatedness on judgements of learning (JOLs). *Memory*, 14(2), 253-261. <https://doi.org/10.1080/09658210500216844>
- Mazancieux, A. (2020). *Towards a unified view of metacognition : Insights from metamemory* [Doctoral thesis]. Université Grenoble Alpes.

- Mazancieux, A., Dinze, C., Souchay, C., & Moulin, C. J. A. (2020). Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence. *Neuroscience of Consciousness*, 2020(1), niaa001. <https://doi.org/10.1093/nc/niaa001>
- Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. A. (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, 149(9), 1788-1799. <https://doi.org/10.1037/xge0000746>
- McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience*, 33(5), 1897-1906. <https://doi.org/10.1523/JNEUROSCI.1890-12.2013>
- McDonough, I. M., Enam, T., Kraemer, K. R., Eakin, D. K., & Kim, M. (2021). Is there more to metamemory? An argument for two specialized monitoring abilities. *Psychonomic Bulletin & Review*, 28(5), 1657-1667. <https://doi.org/10.3758/s13423-021-01930-z>
- McDonough, I. M., & Gallo, D. A. (2012). Illusory expectations can affect retrieval-monitoring accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(2), 391-404. <https://doi.org/10.1037/a0025548>
- McGlynn, S. M., & Kaszniak, A. W. (1991). When metacognition fails : Impaired awareness of deficit in Alzheimer's disease. *Journal of Cognitive Neuroscience*, 3(2), 183-187. <https://doi.org/10.1162/jocn.1991.3.2.183>
- McNair, D. M., & Kahn, R. J. (1983). Self-assessment of cognitive deficit. *Assessment in geriatric psychopharmacology*, 137, 143.
- McWilliams, A., Bibbey, H., Steinbeis, N., David, A. S., & Fleming, S. M. (2022). Age-related decreases in global metacognition are independent of local metacognition and task performance [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/nmhxv>
- Merck, C., Charnallet, A., Auriacombe, S., Belliard, S., Hahn-Barma, V., Kremin, H., Lemesle, B., Mahieux, F., Moreaud, O., Palisson, D. P., Roussel, M., Sellal, F., & Siegwart, H. (2011). La batterie d'évaluation des connaissances sémantiques du GRECO (BECS-GRECO) : Validation et données normatives. *Revue de neuropsychologie*, 3(4), 235. <https://doi.org/10.3917/rne.034.0235>
- Merritt, P., Hirshman, E., Hsu, J., & Berrigan, M. (2005). Metamemory without the memory : Are people aware of midazolam-induced amnesia? *Psychopharmacology*, 177(3), 336-343. <https://doi.org/10.1007/s00213-004-1958-8>
- Metcalfe, J., & Finn, B. (2008a). Familiarity and retrieval processes in delayed judgments of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1084-1097. <https://doi.org/10.1037/a0012580>
- Metcalfe, J., & Finn, B. (2008b). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15(1), 174-179. <https://doi.org/10.3758/PBR.15.1.174>
- Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 19(4), 851-861. <https://doi.org/10.1037/0278-7393.19.4.851>
- Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *The Journal of Neuroscience*, 38(14), 3534-3546. <https://doi.org/10.1523/JNEUROSCI.2360-17.2018>

- Moreaud, O., Belliard, S., Snowden, J. S., Auriacombe, S., Basaglia-Pappas, S., Bernard, F., Bon, L., Boutantin, J., Boutoleau-Brettonnière, C., Charnallet, A., Coutant, E., David, D., Deramecourt, V., Gaestel, Y., Garnier, S., Guichart, E., Hahn-Barma, V., Lebail, B., Lebrun-Givois, C., ... Virat-Brassaud, M.-E. (2008). Semantic dementia : Reflexions of a French working group for diagnostic criteria and constitution of a patient cohort. *Revue Neurologique*, 164(4), 343-353. <https://doi.org/10.1016/j.neurol.2008.02.031>
- Morris, R. G., & Hannesdottir, K. (2004). Loss of 'awareness' in Alzheimer's disease. *Cognitive neuropsychology of Alzheimer's disease*, 275-296.
- Morris, R. G., & Mograbi, D. C. (2013). Anosognosia, autobiographical memory and self knowledge in Alzheimer's disease. *Cortex*, 49(6), 1553-1565.
- Morson, S. M., Moulin, C. J. A., & Souchay, C. (2015). Selective deficits in episodic feeling of knowing in ageing : A novel use of the general knowledge task. *Acta Psychologica*, 157, 85-92. <https://doi.org/10.1016/j.actpsy.2015.02.014>
- Moulin, C. J. A., James, N., Perfect, T. J., & Jones, R. W. (2003). Knowing what you cannot recognise : Further evidence for intact metacognition in Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, 10(1), 74-82. <https://doi.org/10.1076/anec.10.1.74.13456>
- Moulin, C. J. A., Perfect, T. J., & Jones, R. W. (2000). Evidence for intact memory monitoring in Alzheimer's disease : Metamemory sensitivity at encoding. *Neuropsychologia*, 38(9), 1242-1250. [https://doi.org/10.1016/S0028-3932\(00\)00037-3](https://doi.org/10.1016/S0028-3932(00)00037-3)
- Moulin, C. J. A., & Souchay, C. (2014). Epistemic feelings and memory. In T. J. Perfect & D. S. Lindsay, *The SAGE Handbook of Applied Memory* (p. 520-538). SAGE Publications Ltd. <https://doi.org/10.4135/9781446294703.n29>
- Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning : Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, 70, 1-12. <https://doi.org/10.1016/j.jml.2013.09.007>
- Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, 20(2), 378-384. <https://doi.org/10.3758/s13423-012-0343-6>
- Murphy, D. H., Huckins, S. C., Rhodes, M. G., & Castel, A. D. (2022). The effect of perceptual processing fluency and value on metacognition and remembering. *Psychonomic Bulletin & Review*, 29(3), 910-921. <https://doi.org/10.3758/s13423-021-02030-8>
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance : Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1170-1187. <https://doi.org/10.1037/0278-7393.26.5.1170>
- Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults : Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 520-537. <https://doi.org/10.1037/0278-7393.31.3.520>
- Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory : Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 826-837. <https://doi.org/10.1037/0278-7393.29.5.826>

- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P. H., Albert, M., Boone, K., Miller, B. L., Cummings, J., & Benson, D. F. (1998). Frontotemporal lobar degeneration : A consensus on clinical diagnostic criteria. *Neurology*, 51(6), 1546-1554. <https://doi.org/10.1212/WNL.51.6.1546>
- Neely, J. H. (1991). Semantic priming effects in visual word recognition : A selective review of current findings and theories. In *Basic Processes in Reading : Visual Word Recognition* (Vol. 11, p. 264-336).
- Neely, J. H., & Kahan, T. A. (2001). Is semantic activation automatic? A critical re-evaluation. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Éds.), *The nature of remembering : Essays in honor of Robert G. Crowder*. (p. 69-93). American Psychological Association. <https://doi.org/10.1037/10394-005>
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95(1), 109-133.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, 51(2), 102-116. <https://doi.org/10.1037/0003-066X.51.2.102>
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science*, 2(4), 267-271. <https://doi.org/10.1111/j.1467-9280.1991.tb00147.x>
- Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, 113(2), 282-300. <https://doi.org/10.1037/0096-3445.113.2.282>
- Nelson, T. O., McSpadden, M., Fromme, K., & Marlatt, G. A. (1986). Effects of alcohol intoxication on metamemory and on retrieval from long-term memory. *Journal of Experimental Psychology: General*, 115(3), 247-254. <https://doi.org/10.1037/0096-3445.115.3.247>
- Nelson, T. O., & Narens, L. (1990). Metamemory : A Theoretical Framework and New Findings. In *Psychology of Learning and Motivation* (Vol. 26, p. 125-173). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2 : A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516-524. <https://doi.org/10.3758/BF03195598>
- Noble, C. E. (1954). The familiarity-frequency relationship. *Journal of Experimental Psychology*, 47(1), 13-16. <https://doi.org/10.1037/h0060025>
- Norman, E. (2020). Why metacognition is not always helpful. *Frontiers in Psychology*, 11, 1537. <https://doi.org/10.3389/fpsyg.2020.01537>
- Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in psychology. *Review of General Psychology*, 23(4), 403-424. <https://doi.org/10.1177/1089268019883821>
- Öhlschläger, S., & Vö, M. L.-H. (2017). SCEGRAM : An image database for semantic and syntactic inconsistencies in scenes. *Behavior Research Methods*, 49(5), 1780-1791. <https://doi.org/10.3758/s13428-016-0820-3>
- Olney, N. T., Spina, S., & Miller, B. L. (2017). Frontotemporal dementia. *Neurologic Clinics*, 35(2), 339-374. <https://doi.org/10.1016/j.ncl.2017.01.008>

- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, 12(6), 237-241. <https://doi.org/10.1016/j.tics.2008.02.014>
- Ortiz-Tudela, J., Milliken, B., Botta, F., LaPointe, M., & Lupiañez, J. (2017). A cow on the prairie vs. a cow on the street: Long-term consequences of semantic conflict on episodic encoding. *Psychological Research*, 81(6), 1264-1275. <https://doi.org/10.1007/s00426-016-0805-y>
- Packard, P. A., Rodríguez-Fornells, A., Bunzeck, N., Nicolás, B., de Diego-Balaguer, R., & Fuentemilla, L. (2017). Semantic congruence accelerates the onset of the neural signals of successful memory encoding. *The Journal of Neuroscience*, 37(2), 291-301. <https://doi.org/10.1523/JNEUROSCI.1622-16.2016>
- Packard, P. A., Steiger, T. K., Fuentemilla, L., & Bunzeck, N. (2020). Neural oscillations and event-related potentials reveal how semantic congruence drives long-term memory in both young and older humans. *Scientific Reports*, 10(1), 9116. <https://doi.org/10.1038/s41598-020-65872-7>
- Paivio, A. (1990). *Mental Representations*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195066661.001.0001>
- Paivio, A. (2010). Dual coding theory and the mental lexicon. *The Mental Lexicon*, 5(2), 205-230. <https://doi.org/10.1075/ml.5.2.04pai>
- Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1196-1204. <https://doi.org/10.1037/0278-7393.20.5.1196>
- Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and Cognition*, 28, 151-160. <https://doi.org/10.1016/j.concog.2014.06.007>
- Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review*, 15(3), 105-130. <https://doi.org/10.1007/s11065-005-7091-6>
- Pannu, J. K., Kaszniak, A. W., & Rapcsak, S. Z. (2005). Metamemory for faces following frontal lobe damage. *Journal of the International Neuropsychological Society: JINS*, 11(6), 668-676. <https://doi.org/10.1017/S1355617705050873>
- Pappas, B. A., Sunderland, T., Weingartner, H. M., Vitiello, B., Martinson, H., & Putnam, K. (1992). Alzheimer's disease and feeling-of-knowing for knowledge and episodic memory. *Journal of Gerontology*, 47(3), P159-P164. <https://doi.org/10.1093/geronj/47.3.P159>
- Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging*, 7(2), 290-298. <https://doi.org/10.1037/0882-7974.7.2.290>
- Pattamadilok, C., Chanoine, V., Pallier, C., Anton, J.-L., Nazarian, B., Belin, P., & Ziegler, J. C. (2017). Automaticity of phonological and semantic processing during visual word recognition. *NeuroImage*, 149, 244-255. <https://doi.org/10.1016/j.neuroimage.2017.02.003>
- Patterson, K., & Lambon Ralph, M. A. (2016). The hub-and-spoke hypothesis of semantic memory. In *Neurobiology of Language* (p. 765-775). Elsevier. <https://doi.org/10.1016/B978-0-12-407794-2.00061-4>
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976-987. <https://doi.org/10.1038/nrn2277>

- Perfect, T. J., Williams, R. B., & Anderton-brown, C. (1995). Age differences in reported recollective experience are due to encoding effects, not response bias. *Memory*, 3(2), 169-186. <https://doi.org/10.1080/09658219508258964>
- Perrotin, A., Isingrini, M., Souchay, C., Clarys, D., & Taconnat, L. (2006). Episodic feeling-of-knowing accuracy and cued recall in the elderly : Evidence for double dissociation involving executive functioning and processing speed. *Acta Psychologica*, 122(1), 58-73. <https://doi.org/10.1016/j.actpsy.2005.10.003>
- Perrotin, A., Tournelle, L., & Isingrini, M. (2008). Executive functioning and memory as potential mediators of the episodic feeling-of-knowing accuracy. *Brain and Cognition*, 67(1), 76-87. <https://doi.org/10.1016/j.bandc.2007.11.006>
- Piolino, P., Belliard, S., Desgranges, B., Perron, M., & Eustache, F. (2003). Autobiographical memory and autoneotic consciousness in a case of semantic dementia. *Cognitive Neuropsychology*, 20(7), 619-639. <https://doi.org/10.1080/02643290242000899>
- Pliske, R. M., & Mutter, S. A. (1996). Age differences in the accuracy of confidence judgments. *Experimental Aging Research*, 22(2), 199-216. <https://doi.org/10.1080/03610739608254007>
- PRISMA-P Group, Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. <https://doi.org/10.1186/2046-4053-4-1>
- Psychology Software Tools, I. (2016). *E-Prime 3.0. Psychology Software Tools, Inc.*
- Pulvermüller, F. (2013). How neurons make meaning : Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 17(9), 458-470. <https://doi.org/10.1016/j.tics.2013.06.004>
- Putnam, A. L., Deng, W., & DeSoto, K. A. (2022). Confidence ratings are better predictors of future performance than delayed judgments of learning. *Memory*, 30(5), 537-553. <https://doi.org/10.1080/09658211.2022.2026973>
- Rabinowitz, J. C., Ackerman, B. P., Craik, F. I., & Hinchley, J. L. (1982). Aging and metamemory : The roles of relatedness and imagery. *Journal of Gerontology*, 37(6), 688-695.
- Rahnev, D., & Fleming, S. M. (2019). How experimental procedures influence estimates of metacognitive ability. *Neuroscience of Consciousness*, 2019(1), niz010. <https://doi.org/10.1093/nc/niz010>
- Rajaram, S., Hamilton, M., & Bolton, A. (2002). Distinguishing states of awareness from confidence during retrieval : Evidence from amnesia. *Cognitive, Affective, & Behavioral Neuroscience*, 2(3), 227-235.
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M.-L., Rosen, H., ... Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456-2477. <https://doi.org/10.1093/brain/awr179>
- Read, J. D., & Bruce, D. (1982). Longitudinal tracking of difficult memory retrievals. *Cognitive Psychology*, 14(2), 280-300. [https://doi.org/10.1016/0010-0285\(82\)90011-1](https://doi.org/10.1016/0010-0285(82)90011-1)

- Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, 19(1), 90-138. [https://doi.org/10.1016/0010-0285\(87\)90005-3](https://doi.org/10.1016/0010-0285(87)90005-3)
- Reder, L. M. (1988). Strategic control of retrieval strategies. In *Psychology of Learning and Motivation* (Vol. 22, p. 227-259). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60042-0](https://doi.org/10.1016/S0079-7421(08)60042-0)
- Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(3), 435-451. <https://doi.org/10.1037/0278-7393.18.3.435>
- Reggev, N., Zuckerman, M., & Maril, A. (2011). Are all judgments created equal? *Neuropsychologia*, 49(5), 1332-1342. <https://doi.org/10.1016/j.neuropsychologia.2011.01.013>
- Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering : The semantic–episodic distinction. *Trends in Cognitive Sciences*, 23(12), 1041-1057. <https://doi.org/10.1016/j.tics.2019.09.008>
- Renoult, L., & Rugg, M. D. (2020). An historical perspective on Endel Tulving’s episodic-semantic distinction. *Neuropsychologia*, 139, 107366. <https://doi.org/10.1016/j.neuropsychologia.2020.107366>
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information : Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, 137(4), 615-625. <https://doi.org/10.1037/a0013684>
- Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information : Effects on monitoring and control. *Psychonomic Bulletin & Review*, 16(3), 550-554. <https://doi.org/10.3758/PBR.16.3.550>
- Roderer, T., & Roebbers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory : An eye-tracking approach. *Metacognition and Learning*, 5(3), 229-250. <https://doi.org/10.1007/s11409-010-9059-z>
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories : Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803-814. <https://doi.org/10.1037/0278-7393.21.4.803>
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall : A multiple regression analysis. *Psychonomic Bulletin & Review*, 8(3), 385-407. <https://doi.org/10.3758/BF03196177>
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory : A neuropsychological and computational investigation. *Psychological Review*, 111(1), 205-235. <https://doi.org/10.1037/0033-295X.111.1.205>
- Rosen, H. J., Alcantar, O., Zakrzewski, J., Shimamura, A. P., Neuhaus, J., & Miller, B. L. (2014). Metacognition in the behavioral variant of frontotemporal dementia and Alzheimer’s disease. *Neuropsychologia*, 28(3), 436-447. <https://doi.org/10.1037/neu0000012>
- Rosenthal, D. (2000). Consciousness and metacognition. *Metarepresentation: Proceedings of the tenth Vancouver cognitive science conference*, 265-295.
- Rosenthal, D. (2004). Varieties of higher-order theory. *Higher-order theories of consciousness*, 17-44.

- Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition : Evidence from ROC curves. *Journal of Memory and Language*, *43*(1), 67-88. <https://doi.org/10.1006/jmla.1999.2701>
- Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience*, *1*, e17. <https://doi.org/10.1017/pen.2018.16>
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225-237. <https://doi.org/10.3758/PBR.16.2.225>
- Roux-Sibilon, A., & Peyrin, C. (2019, septembre 25). *ScenObjects database*. [osf.io/ur9y5](https://osf.io/ur9y5)
- Rouy, M., Saliou, P., Nalborczyk, L., Pereira, M., Roux, P., & Faivre, N. (2021). Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neuroscience & Biobehavioral Reviews*, *126*, 329-337. <https://doi.org/10.1016/j.neubiorev.2021.03.017>
- Ruijter, N. S., Schoonbrood, A. M. G., Twillert, B., & Hoff, E. I. (2020). Anosognosia in dementia: A review of current assessment instruments. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *12*(1). <https://doi.org/10.1002/dad2.12079>
- Sacher, M., Isingrini, M., & Taconnat, L. (2013). Effects of aging and divided attention on episodic feeling-of-knowing accuracy. *Acta Psychologica*, *144*(2), 258-263. <https://doi.org/10.1016/j.actpsy.2013.07.004>
- Sacher, M., Landré, L., & Taconnat, L. (2015). Age-related differences in episodic feeling-of-knowing arise from differences in memory performance. *Memory*, *23*(2), 119-126. <https://doi.org/10.1080/09658211.2013.870210>
- Sauer, J. D., Palmer, M. A., & Brewer, N. (2019). Pitfalls in using eyewitness confidence to diagnose the accuracy of an individual identification decision. *Psychology, Public Policy, and Law*, *25*(3), 147-165. <https://doi.org/10.1037/law0000203>
- Schacter, D. L. (1983). Feeling of knowing in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 39-54. <https://doi.org/10.1037/0278-7393.9.1.39>
- Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults : The distinctiveness heuristic. *Journal of Memory and Language*, *40*(1), 1-24. <https://doi.org/10.1006/jmla.1998.2611>
- Schacter, D. L., & Worling, J. R. (1985). Attribute information and the feeling-of-knowing. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *39*(3), 467-475. <https://doi.org/10.1037/h0080074>
- Schaper, M. L., & Bayen, U. J. (2021). The metamemory expectancy illusion in source monitoring affects metamemory control and memory. *Cognition*, *206*, 104468. <https://doi.org/10.1016/j.cognition.2020.104468>
- Schaper, M. L., Kuhlmann, B. G., & Bayen, U. J. (2019a). Metamemory expectancy illusion and schema-consistent guessing in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 470-496. <https://doi.org/10.1037/xlm0000602>
- Schaper, M. L., Kuhlmann, B. G., & Bayen, U. J. (2019b). Metacognitive expectancy effects in source monitoring : Beliefs, in-the-moment experiences, or both? *Journal of Memory and Language*, *107*, 95-110. <https://doi.org/10.1016/j.jml.2019.03.009>

- Schnyer, D. M., Nicholls, L., & Verfaellie, M. (2005). The role of VMPC in metamemorial judgments of content retrievability. *Journal of Cognitive Neuroscience*, 17(5), 832-846. <https://doi.org/10.1162/0898929053747694>
- Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., & Kaszniak, A. W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments : Evidence from patients with lesions to frontal cortex. *Neuropsychologia*, 42(7), 957-966. <https://doi.org/10.1016/j.neuropsychologia.2003.11.020>
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26(1/2), 113-125. <https://doi.org/10.1023/A:1003044231033>
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, 4(1), 33-45. <https://doi.org/10.1007/s11409-008-9031-3>
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351-371. <https://doi.org/10.1007/BF02212307>
- Schulman, A. I. (1974). Memory for words recently classified. *Memory & Cognition*, 2(1), 47-52. <https://doi.org/10.3758/BF03197491>
- Schwartz, B. L. (1994). Sources of information in metamemory : Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review*, 1(3), 357-375. <https://doi.org/10.3758/BF03213977>
- Schwartz, B. L. (1999). Sparkling at the end of the tongue : The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, 6(3), 379-393. <https://doi.org/10.3758/BF03210827>
- Schwartz, B. L. (2006). Tip-of-the-tongue states as metacognition. *Metacognition and Learning*, 1(2), 149-158. <https://doi.org/10.1007/s11409-006-9583-z>
- Schwartz, B. L., Boduroglu, A., & Tekcan, A. İ. (2016). Methodological concerns : The feeling-of-knowing task affects resolution. *Metacognition and Learning*, 11(3), 305-316. <https://doi.org/10.1007/s11409-015-9152-4>
- Schwartz, B. L., & Frazier, L. D. (2005). Tip-of-the-tongue states and aging : Contrasting psycholinguistic and metacognitive perspectives. *The Journal of General Psychology*, 132(4), 377-391. <https://doi.org/10.3200/GENP.132.4.377-391>
- Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1074-1083. <https://doi.org/10.1037/0278-7393.18.5.1074>
- Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. *Metacognition: Knowing about knowing*, 93-113.
- Schwartz, B. L., Pillot, M., & Bacon, E. (2014). Contextual information influences the feeling of knowing in episodic memory. *Consciousness and Cognition*, 29, 96-104. <https://doi.org/10.1016/j.concog.2014.08.018>
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568. <https://doi.org/10.1037/0033-295X.96.4.523>
- Seow, T. X. F., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How local and global metacognition shape mental health. *Biological Psychiatry*, 90(7), 436-446. <https://doi.org/10.1016/j.biopsych.2021.05.013>

- Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory : A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(3), 452-460. <https://doi.org/10.1037/0278-7393.12.3.452>
- Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I was so sure! Metacognitive judgments are less accurate given prospectively than retrospectively. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00218>
- Signoret, J.-L. (1991). *Batterie d'efficience mnésique*. Elsevier.
- Simons, J. S., Graham, K. S., Galton, C. J., Patterson, K., & Hodges, J. R. (2001). Semantic knowledge and episodic memory for faces in semantic dementia. *Neuropsychology*, 15(1), 101-114. <https://doi.org/10.1037/0894-4105.15.1.101>
- Simons, J. S., Graham, K. S., & Hodges, J. R. (2002). Perceptual and semantic contributions to episodic memory : Evidence from semantic dementia and Alzheimer's disease. *Journal of Memory and Language*, 47(2), 197-213. [https://doi.org/10.1016/S0749-596X\(02\)00003-7](https://doi.org/10.1016/S0749-596X(02)00003-7)
- Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia : A form of circumscribed cerebral atrophy. *Behavioural Neurology*, 2(3), 167-182. <https://doi.org/10.1155/1989/124043>
- Snowden, J. S., Thompson, J. C., & Neary, D. (2004). Knowledge of famous faces and names in semantic dementia. *Brain*, 127(4), 860-872. <https://doi.org/10.1093/brain/awh099>
- Son, L. K., & Kornell, N. (2008). Research on the allocation of study time : Key studies from 1890 to the present (and beyond). *A handbook of memory and metamemory*, 333-351.
- Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 204-221. <https://doi.org/10.1037/0278-7393.26.1.204>
- Son, L. K., & Metcalfe, J. (2005). Judgments of learning : Evidence for a two-stage process. *Memory & Cognition*, 33(6), 1116-1129. <https://doi.org/10.3758/BF03193217>
- Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and Cognition*, 20(4), 1787-1792. <https://doi.org/10.1016/j.concog.2010.12.011>
- Souchay, C. (2007). Metamemory in Alzheimer's disease. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 43(7), 987-1003.
- Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychologica*, 139(3), 458-464. <https://doi.org/10.1016/j.actpsy.2012.01.007>
- Souchay, C., Isingrini, M., & Espagnet, L. (2000). Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology*, 14(2), 299-309. <https://doi.org/10.1037/0894-4105.14.2.299>
- Souchay, C., Isingrini, M., & Gil, R. (2002). Alzheimer's disease and feeling-of-knowing in episodic memory. *Neuropsychologia*, 40(13), 2386-2396. [https://doi.org/10.1016/S0028-3932\(02\)00075-1](https://doi.org/10.1016/S0028-3932(02)00075-1)

- Souchay, C., Isingrini, M., Pillon, B., & Gil, R. (2003). Metamemory accuracy in Alzheimer's disease and frontotemporal lobe dementia. *Neurocase*, 9(6), 482-492. <https://doi.org/10.1076/neur.9.6.482.29376>
- Souchay, C., Moulin, C. J. A., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: Evidence from feeling-of-knowing and recollection. *Consciousness and Cognition*, 16(4), 769-784. <https://doi.org/10.1016/j.concog.2006.11.002>
- Sterne, J. A. C., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *The Stata Journal*, 4(2), 127-141. <https://doi.org/10.1177/1536867X0400400204>
- Straub, H. R., & Granaas, M. M. (1992). Interaction of instructions with the recall strategy actually used in a paired-associates learning task. *Psychological Reports*, 71(3), 987-993. <https://doi.org/10.2466/pr0.1992.71.3.987>
- Sunderaraman, P., & Cosentino, S. (2017). Integrating the constructs of anosognosia and metacognition: A review of recent findings in dementia. *Current Neurology and Neuroscience Reports*, 17(3), 27. <https://doi.org/10.1007/s11910-017-0734-1>
- Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency: Word familiarity and frequency. *Studia Linguistica*, 65(1), 96-116. <https://doi.org/10.1111/j.1467-9582.2010.01176.x>
- Tauber, S. K., & Rhodes, M. G. (2012a). Multiple bases for young and older adults' judgments of learning in multitrial learning. *Psychology and Aging*, 27(2), 474-483. <https://doi.org/10.1037/a0025246>
- Tauber, S. K., & Rhodes, M. G. (2012b). Measuring memory monitoring with judgements of retention (JORs). *Quarterly Journal of Experimental Psychology*, 65(7), 1376-1396. <https://doi.org/10.1080/17470218.2012.656665>
- Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2011). Context affects feeling-of-knowing accuracy in younger and older adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 37(1), 96-108. <https://doi.org/10.1037/a0021612>
- Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2012). An analysis of the determinants of the feeling of knowing. *Consciousness and Cognition*, 21(4), 1681-1694. <https://doi.org/10.1016/j.concog.2012.09.005>
- Townsend, C. L., & Heit, E. (2011). Judgments of learning and improvement. *Memory & Cognition*, 39(2), 204-216. <https://doi.org/10.3758/s13421-010-0019-2>
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, *Organization of memory* (p. 381-403). Academic Press.
- Tulving, E. (1983). *Elements of episodic memory* (Reprinted). Clarendon Press.
- Tulving, E. (1984). Relations among components and processes of memory. *Behavioral and Brain Sciences*, 7(2), 257-268. <https://doi.org/10.1017/S0140525X00044617>
- Tulving, E. (1985a). How many memory systems are there? *American Psychologist*, 40(4), 385-398. <https://doi.org/10.1037/0003-066X.40.4.385>
- Tulving, E. (1985b). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, 26(1), 1-12. <https://doi.org/10.1037/h0080017>
- Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, 2(3), 67-70. <https://doi.org/10.1111/1467-8721.ep10770899>

- Tulving, E. (1995). Organization of memory : Quo vadis? In *The cognitive neurosciences*. (p. 839-853). The MIT Press.
- Underwood, B. J. (1966). Individual and group predictions of item difficulty for free learning. *Journal of Experimental Psychology*, *71*(5), 673-679. <https://doi.org/10.1037/h0023107>
- Undorf, M., & Bröder, A. (2020). Cue integration in metamemory judgements is strategic. *Quarterly Journal of Experimental Psychology*, *73*(4), 629-642. <https://doi.org/10.1177/1747021819882308>
- Undorf, M., & Bröder, A. (2021). Metamemory for pictures of naturalistic scenes : Assessment of accuracy and cue utilization. *Memory & Cognition*, *49*(7), 1405-1422. <https://doi.org/10.3758/s13421-021-01170-5>
- Undorf, M., & Erdfelder, E. (2015). The relatedness effect on judgments of learning : A closer look at the contribution of processing fluency. *Memory & Cognition*, *43*(4), 647-658. <https://doi.org/10.3758/s13421-014-0479-x>
- Undorf, M., Söllner, A., & Bröder, A. (2018). Simultaneous utilization of multiple cues in judgments of learning. *Memory & Cognition*, *46*(4), 507-519. <https://doi.org/10.3758/s13421-017-0780-6>
- Undorf, M., Zimdahl, M. F., & Bernstein, D. M. (2017). Perceptual fluency contributes to effects of stimulus size on judgments of learning. *Journal of Memory and Language*, *92*, 293-304. <https://doi.org/10.1016/j.jml.2016.07.003>
- Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In *The experience of thinking* (p. 11-32). Psychology Press.
- Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking : A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances*, *2*, 239821281881059. <https://doi.org/10.1177/2398212818810591>
- Valdois, S., Carbonnel, S., Juphard, A., Baciú, M., Ans, B., Peyrin, C., & Segebarth, C. (2006). Polysyllabic pseudo-word processing in reading and lexical decision: Converging evidence from behavioral data, connectionist simulations and functional MRI. *Brain Research*, *1085*(1), 149-162. <https://doi.org/10.1016/j.brainres.2006.02.049>
- van Kesteren, M. T. R., Ruitter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211-219. <https://doi.org/10.1016/j.tins.2012.02.001>
- Vuorre, M., & Metcalfe, J. (2021). Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacognition and Learning*. <https://doi.org/10.1007/s11409-020-09257-1>
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*(4), 635-657. <https://doi.org/10.1080/14640747508400525>
- Warrington, E. K., & Cipolotti, L. (1996). Word comprehension : The distinction between refractory and storage impairments. *Brain*, *119*(2), 611-625. <https://doi.org/10.1093/brain/119.2.611>
- Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, *106*(4), 859-878. <https://doi.org/10.1093/brain/106.4.859>
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology A*, *50*(2), 439-456. <https://doi.org/10.1080/027249897392170>

- Wells, A. (2011). *Metacognitive therapy for anxiety and depression*. Guilford press.
- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272-292.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1235-1253. <https://doi.org/10.1037/0278-7393.19.6.1235>
- Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, *29*(6), 716-732. [https://doi.org/10.1016/0749-596X\(90\)90045-2](https://doi.org/10.1016/0749-596X(90)90045-2)
- Wilson, R. S., Sytsma, J., Barnes, L. L., & Boyle, P. A. (2016). Anosognosia in dementia. *Current Neurology and Neuroscience Reports*, *16*(9), 77. <https://doi.org/10.1007/s11910-016-0684-z>
- Wojcik, D. Z., Moulin, C. J. A., & Souchay, C. (2013). Metamemory in children with autism: Exploring “feeling-of-knowing” in episodic and semantic memory. *Neuropsychology*, *27*(1), 19-27. <https://doi.org/10.1037/a0030526>
- Wojcik, D. Z., & Souchay, C. (2013). Metamemory in children and adolescents with autism spectrum disorder. *OA Autism*, *1*(2). <https://doi.org/10.13172/2052-7810-1-2-595>
- Wojcik, D. Z., Waterman, A. H., Lestié, C., Moulin, C. J. A., & Souchay, C. (2014). Metacognitive judgments-of-learning in adolescents with autism spectrum disorder. *Autism*, *18*(4), 393-408. <https://doi.org/10.1177/1362361313479453>
- Xue, K., Shekhar, M., & Rahnev, D. (2021). Examining the robustness of the relationship between metacognitive efficiency and metacognitive bias. *Consciousness and Cognition*, *95*, 103196. <https://doi.org/10.1016/j.concog.2021.103196>
- Yang, C., Huang, T. S.-T., & Shanks, D. R. (2018). Perceptual fluency affects judgments of learning: The font size effect. *Journal of Memory and Language*, *99*, 99-110. <https://doi.org/10.1016/j.jml.2017.11.005>
- Yee, E. (2019). Abstraction and concepts: When, how, where, what and why? *Language, Cognition and Neuroscience*, *34*(10), 1257-1265. <https://doi.org/10.1080/23273798.2019.1660797>
- Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2014). Semantic memory. In *The Oxford handbook of cognitive neuroscience, Vol. 1: Core topics* (p. 353-374). Oxford University Press.
- Yee, E., Jones, M. N., & McRae, K. (2018). Semantic Memory. In J. T. Wixted (Éd.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (p. 1-38). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119170174.epcn309>
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341-1354. <https://doi.org/10.1037/0278-7393.20.6.1341>
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441-517. <https://doi.org/10.1006/jmla.2002.2864>

- Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*, *41*(2), 229-241. <https://doi.org/10.3758/s13421-012-0255-8>
- Zakay, D., & Tuvia, R. (1998). Choice latency times as determinants of post-decisional confidence. *Acta Psychologica*, *98*(1), 103-115. [https://doi.org/10.1016/S0001-6918\(97\)00037-1](https://doi.org/10.1016/S0001-6918(97)00037-1)
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>