NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

V

La mémoire comme source de connaissances

Kourken Michaelian

La mémoire est parfois citée comme faisant partie de nos sources épistémiques fondamentales (avec la perception, l'inférence, le témoignage et l'introspection). Toutefois, il n'existe pas de liste non controversée des sources épistémiques fondamentales, et il n'est pas certain que la mémoire puisse être qualifiée de source fondamentale en un sens suffisamment strict pour être intéressant[1]. Mais que la mémoire soit, stricto sensu, fondamentale ou non (la question reste ouverte), il est incontestable qu'elle est fondamentale en un sens moins strict, car elle est clairement une source épistémique *essentielle* : si l'on supprimait

1. Si l'on se fie au critère proposé par Audi [2001], une source est fondamentale si elle est capable de rapporter des connaissances sans dépendre de l'opération d'une autre source. Dans la mesure où elle dépend essentiellement de la réception d'informations fournies par d'autres sources, il apparaît que la mémoire ne satisfait pas le critère proposé par Audi. Toutefois, il n'est pas non plus évident qu'il existe une source qui puisse satisfaire ce critère (y compris la perception, exemple le plus plausible d'une source fondamentale) ; il est donc possible que ce critère soit tout simplement trop exigeant pour pouvoir être satisfait. Peut-être la mémoire sera-t-elle néanmoins capable de satisfaire un critère plus modeste de la fondamentalité épistémique.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

notre capacité de se souvenir, on supprimerait nécessairement une grande partie de la connaissance humaine (peut-être même presque toutes nos connaissances), y compris une forme de connaissance qui, apparemment, est unique à l'espèce humaine.

Premièrement, comme plusieurs épistémologues l'ont remarqué, à tout moment, la plupart de nos croyances ne sont pas occurrentes mais seulement dispositionnelles – c'est-à-dire que, suivant une conception naturelle de la croyance dispositionnelle, elles sont stockées dans la mémoire. Deuxièmement, une forme cruciale de la connaissance du passé (probablement propre à l'espèce humaine [Suddendorf et Corballis, 2007]) dépend de la mémoire en un sens particulièrement important : il est possible pour un sujet d'acquérir une connaissance de son propre passé sans dépendre directement de la mémoire (par exemple, au moyen d'un témoignage reçu de la part d'un autre sujet), mais une forme importante de connaissance de son propre passé (la forme de connaissance qu'un sujet peut avoir de son propre passé mais pas de celui d'un autre sujet) dépend essentiellement de la mémoire – en particulier, de la mémoire épisodique ou souvenir.

En vertu d'un certain nombre de tendances convergentes, la mémoire assume également un rôle social de plus en plus important. Plusieurs de ces tendances soulèvent, de manière directe, des questions philosophiques.

Les guerres de mémoire : l'idée des guerres de mémoire fait référence aux controverses juridiques, politiques et scientifiques autour de la réalité des souvenirs retrouvés (par exemple, des souvenirs d'agression sexuelle durant l'enfance) qui ont pris une grande importance, principalement aux États-Unis, et qui se sont éteintes seulement depuis peu [Hacking, 1995]. En termes scientifiques, le conflit a été productif, puisqu'il a encouragé des recherches sur la nature constructive de la mémoire [Schacter 2001], la fiabilité de la mémoire [Loftus, 1996/1979], et les mécanismes qui permettent d'implanter de faux souvenirs – par exemple la théorie du contrôle de la source [Mitchell et Johnson, 2009]. En termes sociaux, par contre, des vies et des familles ont été détruites. Ainsi, les guerres de mémoire renforcent l'importance d'une question proprement épistémologique sur la mémoire : dans quelles conditions a-t-on le droit de se fier à sa propre mémoire ?

Le « memory boom ». Cela fait un bon moment qu'on parle d'un « memory boom », qui se déploie aussi bien dans la culture que dans la recherche scientifique – surtout dans le domaine de l'histoire [Nora, 1984-1992]. D'un point de vue culturel, l'idée fait référence à une nouvelle obsession liée au passé collectif ; d'un point de vue scientifique, elle

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

fait référence à un nouvel intérêt pour les interactions entre l'histoire et la mémoire collective. David Blight [2009] propose d'expliquer l'apparition du « memory boom » par un nombre de facteurs interagissants qui, ensemble, ont augmenté la pertinence de la mémoire comme phénomène culturel et qui ont encouragé les historiens à focaliser leurs efforts sur la construction de souvenirs collectifs ainsi que sur la différence entre la mémoire et l'histoire. Parmi ceux-ci, la parution d'études sur les mentalités collectives, un nouvel intérêt pour l'histoire sociale, le tournant linguistique dans l'histoire intellectuelle et culturelle, les débats sur le constructionnisme social, une attention à l'identité nationale, et la découverte tardive par les historiens (anglophones) du travail de Maurice Halbwachs [1925/1994] sur les « cadres sociaux de la mémoire ». Les détails de l'explication du *timing* mis de côté, les nouveaux travaux sur la mémoire collective soulèvent plusieurs questions philosophiques intéressantes, y compris celle du statut ontologique de la mémoire collective.

La mémoire et le vieillissement. Le vieillissement de la population en Europe et en Amérique du Nord impose de façon urgente une meilleure compréhension des changements (et des régressions) dans le fonctionnement de la mémoire chez les personnes âgées. Cette nouvelle urgence a mené à une explosion de travaux empiriques sur la mémoire et le vieillissement, notamment sur les changements dans le fonctionnement des formes différentes de la mémoire, ou sur les stratégies compensatrices utilisées par les personnes âgées. On ne s'attendait probablement pas à ce que ces recherches soulèvent des défis pour l'épistémologie et la philosophie de l'esprit. Par exemple, les études du rappel collaboratif et de la mémoire transactive chez les couples mariés âgés [Barnier *et al.*, 2008] se croisent de manière intéressante avec les débats sur la cognition distribuée [Hutchins, 1996] et l'hypothèse de l'esprit étendu [Clark et Chalmers, 1998].

La modification de la mémoire. Avec le développement de nouvelles techniques permettant d'intervenir dans le fonctionnement des systèmes mnésiques (par exemple, des interventions pendant la reconsolidation nous donnent la possibilité de transformer des souvenirs traumatiques [Dudai, 2006]), les implications éthiques et politiques de la modification de la mémoire ont acquis une nouvelle importance. Matthews Liao et Anders Sandberg [2008] décrivent un certain nombre de problèmes éthiques soulevés par les technologies de la modification de la mémoire, parmi lesquels celui de la véracité (la modification de la mémoire a le potentiel de transformer nos croyances sur le monde ainsi que sur nous-mêmes), celui des réactions morales appropriées (la modification de la

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

mémoire a le potentiel de transformer nos sentiments par rapport à des événements importants, ce qui peut donner lieu à des réactions morales inappropriées), celui de l'agentivité (la modification de la mémoire peut nous priver de la possibilité de reconnaître notre agentivité par rapport à certains événements), et celui de l'obligation morale (la modification de la mémoire peut nous empêcher de remplir des obligations de nous souvenir, par exemple dans des cas où nos souvenirs servent d'évidences pour les autres) ; néanmoins, ils soutiennent qu'il n'y a pas de barrière éthique générale empêchant l'utilisation de technologies de modification de la mémoire. Plusieurs de ces problèmes éthiques sont liés à des questions épistémiques concernant les implications de la malléabilité de la mémoire sur sa fiabilité, voire le scepticisme russellien par rapport à la connaissance du passé.

L'explosion de la mémoire externe. L'article dans lequel Andy Clark et David Chalmers [1998] ont proposé l'hypothèse de l'esprit étendu a lancé un débat notable dans les cercles de la philosophie de l'esprit, dans lequel l'idée de mémoire externe a joué un rôle important [Michaelian et Sutton, 2013]. Même si la mémoire externe a principalement servi d'exemple dans ce débat, plutôt que de centre d'intérêt à part entière, de plus en plus de recherches en philosophie et dans d'autres disciplines (psychologie, interaction homme-machine) sont directement consacrées aux interactions entre la mémoire biologique et la mémoire externe, ainsi qu'aux effets cognitifs de notre dépendance croissante aux nouvelles formes numériques de mémoire externe[1]. Ces recherches ont souvent des implications sociales directes (par exemple, en ce qui concerne la protection de la vie privée) mais soulèvent également des questions importantes dans le domaine de l'épistémologie et de la philosophie de l'esprit, notamment celle du statut épistémologique des croyances qui dépendent de la mémoire externe.

Malgré l'intérêt du sujet, la situation actuelle dans le domaine de l'épistémologie de la mémoire semble être plus ou moins analogue à celle de l'épistémologie du témoignage telle qu'elle se présentait avant la parution marquante de l'étude de Tony Coady [1992] : la plupart des épistémologues s'accordent sur le fait que la mémoire est d'une importance capitale pour l'épistémologie, tandis que le nombre de spécialistes dans ce domaine reste assez restreint. Ce manque relatif d'activité théorique provient sans doute en partie de facteurs purement sociologiques. Sur le plan théorique, il est probablement dû à l'influence dominante d'un

1. Michaelian [2012a] ; Arango-Muñoz [2013] ; Clowes [2014].

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

modèle trop simpliste de la mémoire, selon lequel la souvenance est simplement un processus destiné à conserver les croyances (et leurs statuts épistémiques). Si ce modèle est juste, il y aurait effectivement assez peu à dire au sujet de la connaissance mnésique ; il n'est donc pas surprenant que, au fur et à mesure que l'on remplace un modèle qui ne prend pas en compte le riche ensemble de théories et de données provenant de la psychologie par un modèle plus adéquat, le niveau d'activité dans le domaine commence à augmenter.

I. Qu'est-ce que la mémoire ?

I.1. Quantité et fidélité

Donner une description directe de la mémoire n'est pas simple, et, historiquement, on a dû faire appel à de nombreuses métaphores pour tenter de la décrire. La psychologie contemporaine, bien évidemment, a tenté de dépasser ce recours aux métaphores, mais une tension durable persiste entre des métaphores incompatibles, y compris dans le domaine de la psychologie.

Comme l'a montré Douwe Draaisma [2000], bien qu'il nous paraisse aujourd'hui naturel de concevoir la mémoire humaine d'après le modèle de la mémoire d'un ordinateur, il y a, en raison de la grande variété de métaphores qui ont été utilisées pour décrire la mémoire, peu de raisons de supposer qu'elle sera la dernière. En effet, d'une certaine façon, la métaphore de l'ordinateur n'est pas particulièrement originale : la mémoire humaine a souvent été comparée à des technologies nouvelles, y compris à des formes de mémoire externe et des moyens de stockage d'informations. Ainsi, la mémoire a été comparée à une tablette de cire, à une bibliothèque, à des archives, à une photographie, à un film, à un hologramme.

D'après ces métaphores, la souvenance apparaît essentiellement comme l'acte de stocker et de récupérer des unités d'information stables. Mais Draaisma décrit également une autre famille de métaphores, dans plusieurs cas plutôt organiques et non technologiques, tels que des bois, des champs, des labyrinthes, des cavernes, des grottes, des puits de mine, les profondeurs de la mer, des palais, des abbayes, et des théâtres, d'après lesquels la souvenance apparaît comme une activité complexe, difficile, et sujette aux échecs. L'existence de ce deuxième ensemble de métaphores est sans doute due à l'utilité limitée de concevoir la mémoire purement en termes de stockage et de récupération ainsi qu'à la nécessité d'une

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

conception qui souligne que la souvenance est un processus dans lequel l'agent reconstruit activement le passé. Ces deux conceptions générales, représentées par les travaux fondamentaux de Hermann Ebbinghaus [1885] et Frederic Bartlett [1932/1995], sont toujours présentes dans la psychologie contemporaine.

Les recherches d'Ebbinghaus (qui ont mené à sa célèbre courbe d'oubli) étaient concentrées sur la capacité de se souvenir de listes de syllabes sans signification. La conception de la mémoire en termes de quantité qui a animé ces recherches, une conception qui voit la mémoire comme une espèce d'entrepôt où des représentations arbitraires peuvent être déposées, a joué un rôle crucial dans la fondation de la psychologie expérimentale et a dominé la recherche sur la mémoire pendant une longue période après Ebbinghaus (avec, par exemple, le paradigme de l'apprentissage de listes [Koriat et Goldsmith, 1996]). Une conception similaire est aujourd'hui dominante en philosophie, dans la mesure où il est souvent implicitement supposé (surtout lorsque la mémoire n'est pas le sujet de recherche principal) que la souvenance est un processus qui, de manière passive, conserve des croyances.

Comme Asher Koriat et Morris Goldsmith [1996] l'ont remarqué, durant la période récente la conception rivale de la mémoire en termes de fidélité, qui est focalisée sur la correspondance entre les contenus mnésiques et la réalité, est devenue de plus en plus influente, ce qui est dû en grande partie à notre meilleure appréhension des erreurs et des distorsions mnésiques[1]. Les recherches de Bartlett lui-même étaient concentrées sur l'influence des schémas sur la forme des contenus mnésiques, et des thèmes similaires figurent de manière importante dans la recherche contemporaine – par exemple, dans des études sur les effets de la conversation sur ce qui est remémoré [Fagin *et al.*, 2013]. Néanmoins, la tension entre la conception en termes de quantité et la conception en termes de fidélité perdure, et il semble peu vraisemblable qu'elle se résolve dans un avenir proche, car, bien qu'il soit difficile d'intégrer les deux conceptions au sein d'un seul modèle, toutes les deux nous révèlent un aspect important du fonctionnement de la mémoire.

I.2. Les formes de la mémoire

La conception en termes de quantité et la conception en termes de fidélité représentent des tentatives concurrentes pour dire, en termes très

1. Pour une recension accessible, voir Schacter [2001].

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

généraux, comment la mémoire fait ce qu'elle fait. Bien que le désaccord sur le fait de savoir *comment la mémoire fait* ce qu'elle fait ne soit de toute évidence pas en train de se dissiper, un accord général existe, en philosophie ainsi qu'en psychologie, sur *ce que la mémoire fait*. Selon ce consensus, la mémoire n'est pas unitaire : les philosophes ont décrit une variété de formes différentes de la mémoire, et celles-ci correspondent assez bien aux formes distinguées par les psychologues.

I.2.1. L'approche philosophique : mémoire-souvenir et mémoire-habitude

Comme l'a remarqué John Sutton [2010a], les philosophes ont (en utilisant des terminologies différentes) distingué entre une forme de mémoire qui correspond approximativement au *savoir-que* et une forme qui correspond approximativement au *savoir-comment*. Par exemple, Henri Bergson [1896] et Bertrand Russell [1921] distinguent la *mémoire-souvenir* de la *mémoire-habitude*. Au sein de la forme de mémoire qui correspond au savoir-que, les philosophes ont soutenu une distinction entre une forme de mémoire qui correspond à la connaissance factuelle générale et une forme de mémoire qui correspond à la connaissance par un sujet de son passé personnel. Par exemple, Charlie Dunbar Broad [1925] distingue *la mémoire propositionnelle* de *la mémoire-souvenir* à proprement parler.

Ces distinctions philosophiques sont fondées principalement sur des différences grammaticales. On dit, par exemple, que je peux me souvenir de *comment* faire quelque chose, que je peux me souvenir *que* quelque chose est le cas, ou que je peux me souvenir *de* quelque chose, ce qui suggère l'existence de trois formes (au moins) de mémoire avec des types d'objets ou des types de contenus différents. On verra ci-dessous, cependant, que les distinctions philosophiques correspondent de façon approximative aux distinctions établies en psychologie, où le débat portait moins sur la nature des objets des différentes formes de mémoire que sur la nature des mécanismes psychologiques qui donnent lieu à ces formes.

I.2.2. L'approche psychologique : mémoire déclarative, mémoire épisodique, mémoire sémantique

Tandis que les philosophes ont tendance à se fier à des critères grammaticaux afin de distinguer entre des formes différentes de mémoire, les psychologues distinguent souvent des formes différentes de mémoire en vertu de l'hypothèse de systèmes mnésiques multiples. Selon celle-ci, au lieu d'une faculté unitaire de souvenance, il y a plusieurs systèmes

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

mnésiques distincts qui, en fonctionnant séparément ou ensemble, donnent lieu aux diverses formes d'apprentissage et de souvenance. Une possibilité est de concevoir les systèmes distingués par l'hypothèse des systèmes mnésiques multiples en s'appuyant sur une autre hypothèse, celle dite des trois niveaux [Marr, 1982] ; de les concevoir, en d'autres termes, comme des entités distinctes situées au niveau computationnel, au niveau algorithmique, et au niveau implémentationnel [Michaelian, 2011b]. Il semble que Daniel Schacter et Endel Tulving [1994], par exemple, acceptent une telle conception, car ils soutiennent que les systèmes mnésiques doivent être définis selon « la nature de leurs règles d'opération, le type d'informations ou de contenus, et les voies neurales des mécanismes qui les suscitent ». Bien que l'hypothèse des systèmes mnésiques multiples reste populaire, elle ne fait toutefois pas consensus [Foster et Jelicic, 1999].

En dépit d'un désaccord sur l'idée des systèmes mnésiques multiples, les psychologues s'accordent à distinguer la mémoire *déclarative* et la mémoire *non déclarative*. La mémoire non déclarative, qui correspond (très) approximativement à la mémoire-habitude décrite par les philosophes, constitue un ensemble de phénomènes variés. La seule chose que l'on peut probablement affirmer de toutes les formes de mémoire non déclaratives est que, tandis que la mémoire déclarative implique le stockage et la récupération de représentations, la mémoire non déclarative ne peut pas facilement être décrite en de tels termes, car il s'agit plutôt de modifications de comportements et de réponses aux stimuli (par exemple : conditionnement, amorçage). Par conséquent, les concepts d'encodage, de stockage et de récupération, qui nous permettent de définir la mémoire déclarative, ont très peu d'utilité dans le cadre de la définition de la mémoire non déclarative. En fait, on peut soutenir qu'une définition de la mémoire non déclarative en termes de stockage d'informations est simplement erronée [Michaelian, 2011b].

La mémoire déclarative, en revanche, qui dépend du lobe temporal médian, y compris l'hippocampe, ainsi que d'un lieu de stockage spécialisé dans le néocortex, implique des processus d'encodage, de stockage, et de récupération, et concerne le rappel conscient d'informations. Au sein de la mémoire déclarative, on distingue entre la mémoire *épisodique,* qui correspond approximativement à la mémoire-souvenir décrite par les philosophes et qui concerne le rappel d'expériences personnelles, et la mémoire *sémantique,* qui correspond approximativement à la mémoire propositionnelle décrite par les philosophes et qui concerne le rappel de connaissances générales. La mémoire épisodique, contrairement à la

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

mémoire sémantique, dépend des lobes frontaux et implique la conscience autonoétique [Tulving, 2002].

D'ordinaire, les philosophes prêtent une attention particulière à la mémoire épisodique/mémoire souvenir, et, de façon similaire, cette forme de mémoire constitue également un centre d'intérêt important pour la recherche psychologique actuelle, où elle est de plus en plus vue comme un aspect d'une capacité constructive ou simulationnelle plus générale [Klein, 2012]. La plupart des chercheurs distinguent la mémoire épisodique à proprement parler de la mémoire de type épisodique (mémoire « quoi-où-quand ») possédée par certains animaux. Contrairement à la mémoire de type épisodique animale, la mémoire épisodique à proprement parler implique un caractère phénoménal distinctif, la conscience autonoétique – le sujet ne récupère pas simplement une représentation de l'événement, mais a également l'impression subjective de le ré-expérimenter. La mémoire sémantique, par contre, n'implique pas ce caractère phénoménal spécial. Selon une thèse portant sur la nature de la mémoire épisodique qui devient de plus en plus influente, la capacité humaine de se souvenir du passé est étroitement liée à notre capacité d'imaginer l'avenir – la souvenance du passé et l'imagination de l'avenir sont en fait deux aspects d'une capacité générale permettant le voyage mental dans le temps, ce qui, comme l'affirment Thomas Suddendorf et Michael Corballis [2007, p. 299], « nous permet non seulement de retourner dans le temps, mais aussi de prévoir, planifier, et façonner quasiment n'importe quel événement à venir ».

Il y a des raisons évolutionnaires de supposer que l'imagination d'événements à venir (accompagnée de conscience autonoétique) est la fonction principale du système de voyage mental dans le temps [Boyer, 2008]. En fait, les chercheurs ont tendance à voir la souvenance épisodique comme une fonction spécifique d'un système dédié non seulement au voyage mental dans l'avenir et le passé mais surtout à une gamme plus large de fonctions simulationnelles. Une telle approche a été développée récemment par Karen Shanton et Alvin Goldman [2010], en mobilisant la théorie simulationniste de la théorie de l'esprit. Leur suggestion essentielle énonce que, lors de la souvenance, le sujet simule ses propres états mentaux passés, et donc que la simulation est au fond une forme d'imagination épisodique. Des approches similaires de la mémoire épisodique en termes non seulement de voyage mental dans le temps mais également d'autres formes imaginatives de construction d'expériences possibles ont été proposées par d'autres chercheurs. Ainsi, Randy Buckner et Daniel Carroll remarquent que les preuves s'accumulent en faveur de l'idée selon

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

laquelle la mémoire épisodique, la prospection (le voyage mental vers l'avenir), la théorie de l'esprit, et des formes de navigation, sont toutes des produits du même réseau neuronal. Ils proposent une hypothèse selon laquelle la capacité commune dans ce cas est la projection de soi, « la capacité de changer sa perspective du présent immédiat pour d'autres perspectives » [Buckner et Carroll, 2007, p. 49]. Une approche alternative serait ici l'hypothèse de construction de scènes proposée par Demis Hassabis et Eleanor Maguire [2007, p. 299], qui soutiennent que la mémoire épisodique est une parmi plusieurs formes de construction de scènes – « le processus mental de produire et maintenir une scène ou un événement complexe et cohérent » – suscitées par le même réseau neuronal, y compris non seulement la pensée épisodique de l'avenir mais aussi la navigation, la théorie de l'esprit, le *mind-wandering,* et l'imagination d'expériences fictives. Dans toutes ces approches, la mémoire épisodique n'est pas vue comme un système dédié à la conservation d'enregistrements exacts des expériences précédentes du sujet mais plutôt comme une activité d'un système plus général de construction ou de simulation dédié à la production de représentations d'une gamme d'expériences possibles.

La conception contemporaine de la mémoire épisodique comme processus constructif est en désaccord avec la conception de la mémoire comme processus entièrement préservateur qui est adoptée (souvent implicitement) par plusieurs approches philosophiques historiques et contemporaines. Par exemple, Sven Bernecker [2008, p. 145] distingue une conception préservatrice (le « modèle de la photocopieuse ») chez saint Augustin, qui compare la mémoire à une grande salle, un entrepôt, et une caverne, conception selon laquelle les images de choses perçues sont stockées et récupérées en cas de nécessité. Il la reconnaît également dans l'approche de la mémoire adoptée par les empiristes, tels que David Hume [1739] et John Locke [1689]. Plusieurs conceptions contemporaines se laissent également influencer par le modèle de la photocopieuse. Par exemple, la théorie de la mémoire proposée par Bernecker lui-même est compatible uniquement avec des transformations de contenu très limitées dans la souvenance. Tyler Burge [1993] propose une forme purement préservatrice de mémoire. Et Jérôme Dokic [2001a] juge la mémoire sémantique purement préservatrice, et considère que des transformations de contenu très limitées interviennent dans la mémoire épisodique. On peut effectivement affirmer que n'importe quelle analyse de la souvenance qui intègre une stricte condition de représentation antérieure doit nécessairement accepter une version

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

du modèle de la photocopieuse, et doit ainsi risquer de se rendre incompatible avec la vue contemporaine de la mémoire dite constructive.

II. Théories de la souvenance

Il y a relativement peu de recherches sur l'analyse de la mémoire non déclarative[1]. Les philosophes ont plutôt concentré leurs efforts sur la mémoire déclarative, dans une tentative de découvrir, grosso modo, des conditions nécessaires et suffisantes pour des affirmations de la forme :

S se souvient de P/que P.

où P est une proposition (dans le cas de la mémoire sémantique) ou un événement (dans le cas de la mémoire épisodique).

Plus ou moins toutes les théories intègrent une condition de représentation actuelle, grossièrement :

(1) Actuellement, S se représente P.

Comme « souvenir » dans le langage ordinaire se comporte souvent comme un verbe factif – le fait que S se souvienne que P implique que P, tout comme le fait que S sache que P implique que P – la plupart des théories intègrent aussi une condition de vérité :

(2) P est vrai/a eu lieu.

Au-delà de ces deux conditions, il y a peu d'accord entre les théories.

II.1. Théories empiristes

Une suggestion très naturelle, adoptée par la plupart des théoriciens, est d'intégrer une condition de représentation antérieure :

(3) Précédemment, S se représentait P.

1. Ainsi, d'ailleurs, que sur la relation entre la mémoire déclarative et la mémoire non déclarative, contrairement à l'intérêt récent porté à la relation entre le savoir-que et le savoir-comment.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

En fait, on peut dire que la théorie empiriste de la mémoire, qui était influente jusqu'au début du xx^e siècle mais qui est beaucoup moins populaire depuis, consiste essentiellement dans ces trois conditions, auxquelles s'ajoute un critère de vivacité, inclus dans le but de faire la distinction entre la mémoire et l'imagination :

(4) La représentation actuelle de S est suffisamment détaillée.

Hume [1739], en particulier, soutenait que les idées de la mémoire peuvent être distinguées des idées de l'imagination par leur plus grande vivacité. Selon cette approche empiriste, se souvenir consiste plus ou moins à répéter ou à rejouer une expérience précédente.

Il faut remarquer que le critère de vivacité pose un certain nombre de problèmes ; en particulier, il paraît évident que les représentations produites par la mémoire ne sont pas toujours plus détaillées que les représentations produites par l'imagination. (Que l'on considère des cas de ce qu'on peut appeler la construction itérée : se souvenir d'avoir imaginé P, se souvenir de s'être souvenu de P, et ainsi de suite. Comme Bernecker [2008] le remarque, il est peu probable que la représentation impliquée dans la souvenance d'avoir imaginé P soit plus détaillée que la représentation impliquée dans l'imagination de P.) Peut-être est-ce pour cette raison que les autres théories de la mémoire n'intègrent généralement pas le critère de vivacité.

Si l'on ne tient pas compte des problèmes du critère de vivacité, l'approche empiriste de la mémoire est d'une simplicité et d'un naturel attirants. Cependant, l'article classique de Charles B. Martin et Max Deutscher [1966] notamment, qui développe la théorie causale de la mémoire, a soutenu d'une manière convaincante qu'elle est intenable. Ils décrivent le contre-exemple suivant : un sujet observe un événement et forme une croyance que tel et tel événement ont eu lieu ; la croyance est stockée dans sa mémoire ; plus tard, sa mémoire est perdue (à cause d'une lésion cérébrale) ; après avoir perdu la mémoire, le sujet est hypnotisé ; l'hypnotiseur lui implante une croyance que tel et tel événement ont eu lieu ; la nouvelle croyance est indifférenciable de la croyance initiale. Comme la condition de représentation antérieure et le critère de vivacité sont satisfaits dans ce cas, la théorie empiriste implique que le sujet se souvient que l'événement a eu lieu ; mais il est évident que ce n'est pas réellement le cas.

Martin et Deutscher suggèrent que la théorie empiriste implique une classification incorrecte dans de tels cas parce qu'elle n'intègre pas de

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

condition causale. Mais une autre option que la théorie empiriste, à savoir, la théorie épistémique, est disponible, et celle-ci intègre une condition de rétention au lieu d'une condition causale. Il convient d'examiner d'abord la théorie épistémique, avant de revenir à la théorie causale de Martin et Deutscher.

II.2. Théories épistémiques

Les théories épistémiques voient la mémoire comme de la « connaissance retenue », et complètent les conditions 1 à 3 avec une condition de rétention ou une autre condition. Si l'on suit la présentation de Bernecker [2008], les théories épistémiques peuvent être classées en deux groupes : celles qui intègrent une condition de rétention pure, telle que 5, et celles qui intègrent une condition de rétention évidentielle, telle que 6.

(5) S sait que P et a su que P sans interruption depuis qu'il l'a appris.

(6) Les raisons actuelles de S pour accepter P sont les mêmes que ses raisons précédentes pour l'accepter.

Théories de rétention pure. Pour des exemples de conditions de rétention pure, nous pouvons considérer celles proposées par Gilbert Ryle [1949], selon qui la souvenance implique que S n'a pas oublié que P ; Stanley Munsat [1967], selon qui la souvenance implique qu'il faut que S ait toujours su que P entre le moment d'apprentissage et le moment de la présumée souvenance, et Eddie Zemach [1968], selon qui la souvenance implique qu'il faut que S croie qu'il savait précédemment que P.

Comme Bernecker le remarque, des problèmes se posent pour chacune de ces versions de la condition de rétention pure. Les conditions proposées par Ryle et Munsat sont trop fortes, car des cas existent dans lesquels quelqu'un se souvient de quelque chose après l'avoir oublié. En un sens, la condition proposée par Zemach est également trop forte, car quelqu'un peut se souvenir de quelque chose sans croire qu'il le savait précédemment ; en un autre sens, elle est trop faible, car elle classifie certains cas de réapprentissage comme des cas de souvenance (quand S apprend que P, l'oublie complètement, le réapprend, puis croit qu'il le savait précédemment, l'analyse de Zemach implique que S se souvient que P).

Outre ces problèmes, il y a également un problème général pour toute condition de rétention pure, remarqué par Deutscher [1989] et repris par

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

Bernecker : les théories elles-mêmes ne disent pas comment les connaissances sont retenues, et il n'y a pas d'explication plausible de ce processus qui n'implique pas une chaîne causale du genre décrit par la théorie causale de la mémoire ; ce qui suggère que la théorie épistémique, quand elle est expliquée clairement, finit par rejoindre la théorie causale.

Théories de rétention évidentielle. Les théories de rétention évidentielle, au lieu d'affirmer directement qu'il faut que le sujet ait retenu la connaissance, affirment que ses raisons actuelles en faveur de la croyance en question doivent être identiques à ses raisons précédentes. Bernecker localise la première version de cette approche chez Andrew Naylor [1971].

Bernecker soutient que les théories de rétention évidentielle sont confrontées à un certain nombre de problèmes techniques qui peuvent sans doute être surmontés. Par exemple, une condition de rétention évidentielle simple est violée lorsque l'agent acquiert des raisons supplémentaires pour sa croyance. Ce problème peut être évité si l'on exige seulement que les raisons actuelles pour la croyance soient un sur-ensemble des raisons précédentes [Naylor, 1971]. Mais même à supposer que tous les problèmes techniques de ce genre puissent être évités, les théories de rétention évidentielle se trouvent face à un problème plus fondamental : « Il est possible que S sache que P à t_1, se souvienne à t_2 de tout ce qu'il savait à t_1, et ne sache quand même pas que P à t_2 – même s'il continue à croire que P, et P continue d'être vrai » [Bernecker, 2008, p. 29]. La raison de cela est que, dans l'intervalle entre t_1 et t_2, il est possible que S ait rencontré des alternatives pertinentes qu'il n'est pas capable d'éliminer ; et s'il n'est pas capable d'éliminer ces alternatives, il ne sait plus que P, en dépit du fait qu'il se souvienne que P. Autrement dit, « la mémoire est capable de transformer la connaissance en simple croyance vraie » [Bernecker, 2008, p. 29]. La souvenance n'implique pas la connaissance. Si ce raisonnement est correct, il montre qu'aucune théorie épistémique n'est tenable. De façon analogue, Jennifer Lackey [2005] a soutenu que la mémoire est capable de produire de nouvelles connaissances. Le raisonnement de Lackey a des implications similaires pour la théorie épistémique : si la mémoire est capable de produire de nouvelles connaissances, elle ne peut clairement pas être identifiée comme de la connaissance retenue.

II.3. Théories causales

Pour toutes ces raisons, la théorie causale a été plus populaire dernièrement que la théorie épistémique.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

La théorie causale classique. On a déjà vu que Martin et Deutscher affirment que la condition de représentation antérieure doit être complétée par une condition causale, semblable à 7.

(7) La représentation actuelle de S est causée par sa représentation antérieure.

Mais comme ils le reconnaissent eux-mêmes, il n'est pas suffisant de simplement exiger qu'il y ait une connexion causale d'un certain type entre la représentation actuelle et la représentation antérieure, puisqu'on se trouve face au problème des chaînes causales déviantes (des problèmes similaires existent dans d'autres domaines, par exemple, dans l'analyse de la perception) : certaines chaînes causales ne sont pas compatibles avec la souvenance, et l'on a par conséquent besoin de fournir un critère afin de pouvoir décider à quel moment la représentation actuelle se trouve liée à la représentation antérieure par une chaîne causale appropriée.

Martin et Deutscher, par exemple, décrivent le cas suivant : un sujet observe un événement et forme une croyance que tel et tel événement ont eu lieu ; la croyance est stockée dans sa mémoire ; plus tard, le souvenir est perdu (à cause d'une lésion cérébrale) ; mais entre-temps le sujet a décrit l'événement à un ami ; une fois que le sujet a perdu son souvenir, l'ami lui répète sa description de l'événement ; le temps passe et le sujet oublie qu'il a entendu la répétition de la description – maintenant, il croit seulement que tel et tel événement ont eu lieu. Dans ce type de cas, on est en présence d'une chaîne causale qui n'est évidemment pas de la bonne sorte pour permettre la souvenance (épisodique).

Afin de résoudre ce problème, Martin et Deutscher exigent que la chaîne causale passe par une trace mnésique :

(8) La chaîne causale qui relie la représentation actuelle de S à sa représentation antérieure passe par une trace mnésique.

Ce genre de suggestion a été assez controversé, car il semble impliquer que certaines suppositions neuroscientifiques sont intégrées dans notre concept de la souvenance[1].

1. Voir Rosen [1975] pour une tentative de développer un concept purement « logique » des traces mnésiques qui évite ce problème.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

Si notre but est de développer une théorie de la mémoire elle-même, plutôt qu'une théorie de notre concept de la mémoire, l'objection ne nous dérangera pas. Mais Zemach [1983] a proposé une objection quelque peu similaire, qui pose un problème même si la théorie est envisagée comme une théorie de la mémoire elle-même ; l'idée centrale est celle selon laquelle il ne faut pas intégrer une référence à l'existence de traces mnésiques dans une théorie philosophique de la mémoire, parce que la philosophie ne doit pas dicter à la science empirique ce qu'elle doit découvrir dans le cadre du fonctionnement de la mémoire. Cette objection, par contre, dépend de la supposition de Martin et Deutscher selon laquelle les traces mnésiques sont des analogues structurels d'expériences, ce qui exclut des traces distribuées et superpositionnelles du genre décrit par les modèles connexionnistes de la mémoire [Sutton, 1998], qui correspondent mieux à la description du fonctionnement reconstructif de la mémoire issue de la recherche empirique. Le partisan de la théorie causale peut donc éviter l'objection en supposant que les traces pertinentes sont distribuées.

Même si l'on met cette objection de côté, il n'est pas suffisant, comme Martin et Deutscher le reconnaissent, d'exiger que la causation passe par une trace mnésique. Dans l'exemple ci-dessus, la causation passe initialement par une trace, mais le sujet ne se souvient pas malgré tout. De plus, une telle position rencontre des problèmes concernant l'indication : parfois, le sujet ne peut pas se souvenir sans recevoir une indication, et même, dans certains cas, sans recevoir une indication qui contient le contenu entier de ce dont il se souvient ; comment peut-on distinguer les cas où le sujet accepte tout simplement l'indication des cas où, grâce à l'indication, il se souvient réellement ? La solution proposée par Martin et Deutscher consiste essentiellement dans le fait d'exiger que la trace soit opérationnelle dans la production de la présumée représentation mnésique. Être « opérationnel dans » (par opposition à être « opérationnel pour ») est un concept technique introduit par Martin et Deutscher, mais il s'agit essentiellement de la condition suivante :

(9) La trace doit faire une contribution causale à l'instant de la production de la représentation actuelle de S.

La théorie causale de la mémoire constructive. La théorie causale, telle qu'elle a été développée par Martin et Deutscher, a eu un succès impressionnant. Cependant, la théorie causale accepte toujours le modèle de la photocopieuse, puisqu'elle suppose que le fait de se souvenir est

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

essentiellement une question de stockage et de récupération d'une copie d'une représentation. C'est ce qu'implique la conjonction de la condition de représentation actuelle et la condition de représentation antérieure. La version de la théorie causale proposée par Bernecker [2008] permet des transformations de contenu limitées, comme par exemple la généralisation existentielle, mais ne permet pas des transformations qui ajoutent du nouveau contenu ; ainsi, Bernecker lui aussi accepte l'essentiel du modèle de la photocopieuse.

Les travaux sur le caractère constructif de la souvenance cités ci-dessus montrent clairement que, bien que la souvenance fasse appel à des informations stockées, la validité du modèle de la photocopieuse reste très limitée : la mémoire produit régulièrement de nouvelles représentations qui ne correspondent pas exactement aux représentations antérieures, en partie parce qu'elles incluent des informations qui vont au-delà des informations contenues dans les représentations antérieures. En effet, énormément de recherches psychologiques portent sur les distorsions mnésiques produites par cette particularité de la mémoire. Loftus et ses collègues ont ainsi montré que, dans certaines circonstances, il est possible d'implanter des souvenirs d'événements entièrement fictifs [Loftus, 1979/1996].

Une réaction possible face à ces résultats consiste à abandonner entièrement la condition de représentation antérieure, auquel cas il n'est plus nécessaire que la représentation actuelle corresponde à une représentation antérieure. Une telle approche est largement cohérente avec la perspective du système de construction décrit ci-dessus, dans la mesure où elle fait paraître la souvenance du passé essentiellement comme le symétrique de l'imagination de l'avenir : dans les deux cas, le système de construction fait appel à des information stockées afin de simuler une expérience, et la différence entre la souvenance et la pensée épisodique de l'avenir consiste ici seulement dans la direction du sens du temps subjectif impliqué.

Cette approche, cependant, requiert que l'on nie l'existence d'une quelconque différence entre la souvenance du passé et l'imagination du passé. Bien qu'il soit possible qu'il n'y ait finalement pas de différence entre ces deux processus, il est plus conservateur, pour l'instant, de préférer une position qui permet la possibilité d'une telle différence. Il semble en effet que des souvenirs qui n'ont pas la moindre correspondance avec l'expérience antérieure du sujet soient des occurrences peu fréquentes (même s'il est important de remarquer qu'il s'agit d'occurrences ordinaires). Le cas typique concerne plutôt des souvenirs qui s'éloignent dans une

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

certaine mesure de l'expérience originale, tout en gardant un niveau plus ou moins acceptable de similarité globale.

Considérons l'étude d'Ulric Neisser [1981] de la mémoire de John Dean, dans laquelle Neisser découvre que, bien que Dean ait eu tendance à exagérer son propre rôle dans les événements racontés, entre autres distorsions au niveau des détails, le sens général de ses souvenirs était juste. Ou considérons la recherche sur les changements de perspective dans la souvenance [Sutton, 2010b] : dans la plupart des cas, le sujet peut se souvenir d'une expérience soit du point de vue du champ, soit du point de vue de l'observateur (externe, le sujet se voyant éventuellement lui-même dans la scène) ; évidemment, un souvenir d'observateur ne correspond pas exactement à l'expérience antérieure du sujet, mais il peut néanmoins représenter, de manière juste, les aspects les plus importants de cette expérience. Un autre exemple : il y a beaucoup de travaux sur les distorsions dans la mémoire autobiographique, par exemple, sur la tendance à déformer les souvenirs de son moi passé de sorte que ce dernier paraisse pire que le moi actuel, en accord avec le désir de s'améliorer avec le temps [Ross et Wilson, 2003] ; bien que de telles distorsions soient répandues, elles ne suggèrent pas que nos souvenirs portant sur les plans généraux de nos propres vies contiennent des inexactitudes importantes. Un dernier exemple : des recherches sur les distorsions dans les souvenirs flash (les souvenirs particulièrement détaillés d'événements de grande importance) montrent que ces souvenirs, en dépit de leur vivacité phénoménale, sont aussi sujets à des distorsions variées [Talarico et Rubin, 2007] ; mais, là encore, les erreurs ont tendance à s'accumuler dans les questions de détail, plutôt que dans les aspects centraux des événements.

Ces exemples suggèrent que ce dont on a besoin est une approche de la souvenance qui admette un certain niveau de correspondance entre la représentation actuelle du sujet et une représentation précédente, sans requérir une correspondance stricte. On peut donc surmonter le problème posé par le caractère constructif de la souvenance en remplaçant la condition de représentation précédente stricte par une condition moins rigide :

> (3*) Précédemment, S se représentait P, ou quelque chose de suffisamment similaire à P.

Une particularité malheureuse de cette nouvelle condition est qu'elle introduit beaucoup d'ambiguïté au sein de la théorie : quel niveau de similarité entre la représentation actuelle et la représentation antérieure

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

est-il nécessaire de posséder avant de pouvoir dire avec certitude que le sujet se souvient ? Il est peut-être possible de surmonter cette difficulté en supposant que la mémoire épisodique est produite par un système autre que celui responsable de l'imagination ou de la simulation, ou en supposant que, bien que la mémoire épisodique soit produite par un système plus général de construction ou de simulation, la souvenance est une fonction spécifique de ce système, distincte de l'imagination (y compris l'imagination du passé personnel) :

> (10) La chaîne causale qui relie la représentation actuelle de S à sa représentation antérieure passe par un système qui fait appel à la trace mnésique dans le but de construire une représentation correspondant à la représentation antérieure de S.

Si l'on adopte cette stratégie, on permet la compatibilité de la souvenance avec un niveau élevé de construction et cela même avec un niveau très bas de similarité entre la représentation antérieure et la représentation actuelle : pourvu que le système pertinent fasse appel à des informations stockées dans le but de construire une représentation correspondant (dans certaines limites) à une représentation antérieure, on peut dire que le sujet se souvient.

En même temps, le fait que les distorsions soient répandues dans la mémoire suggère que la théorie de la mémoire ne devrait pas intégrer une condition de vérité trop stricte. Évidemment, la justesse joue un rôle en déterminant si un sujet se souvient ou non (un sujet avec un système mnésique qui, à cause d'un mauvais fonctionnement, produit des représentations qui n'ont pas la moindre correspondance avec ses expériences antérieures n'arrive pas à se souvenir), mais, si l'on considère l'étendue des distorsions au niveau des détails, une condition de vérité trop stricte impliquera que la souvenance réussie est extrêmement rare. Ainsi, la condition de vérité 2 devrait être remplacée par une condition assez proche :

> (2*) P est approximativement correct/quelque chose de similaire à P a eu lieu.

Évidemment, cette stratégie présuppose que nous ayons des concepts acceptables de vérité approximative et de similarité entre événements.

Quoiqu'il reste encore beaucoup à faire, il semble donc que la théorie suivante (similaire à la théorie causale de la mémoire constructive développée par Michaelian [2011b]) soit prometteuse et compatible avec nos

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

connaissances scientifiques les plus avancées : S se souvient de P/que P si et seulement si :

(1) Actuellement, S se représente P ;

(2*) P est approximativement correct/quelque chose de similaire à P a eu lieu ;

(3*) Précédemment, S se représentait P, ou quelque chose de suffisamment similaire à P ;

(7) La représentation actuelle de S est causée par sa représentation antérieure ;

(8) La chaîne causale qui relie la représentation actuelle de S à sa représentation précédente passe par une trace mnésique ;

(9) La trace doit compter causalement lors de la production de la représentation actuelle de S ;

(10) La chaîne causale qui relie la représentation actuelle de S à sa représentation antérieure passe par un système qui fait appel à la trace mnésique dans le but de construire une représentation correspondant à la représentation antérieure de S.

Implications épistémologiques de la théorie causale de la mémoire constructive. Lackey [2005] a argumenté contre la conception préservationniste de la justification mnésique, peut-être la conception la plus naturelle, en soutenant une forme modérée de générationnisme, selon laquelle la mémoire est capable d'engendrer de la justification en produisant une nouvelle croyance avec un contenu non endossé auparavant. L'essentiel de l'argument s'énonce ainsi : lorsqu'un sujet acquiert initialement une information, cette dernière peut être stockée dans sa mémoire sans que le sujet forme d'abord une croyance ayant l'information comme contenu ; une récupération subséquente de l'information peut mener à la formation d'une telle croyance ; si les processus cognitifs pertinents sont suffisamment fiables, alors cette croyance sera justifiée. Thomas Senor [2007] a argumenté contre le générationnisme modéré[1], mais, si la

1. Voir Lackey [2007a] pour sa propre réponse.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

théorie causale de la mémoire constructive est plus ou moins correcte, le générationnisme modéré semble alors être trop faible : selon la théorie causale de la mémoire constructive, la mémoire est capable d'engendrer non seulement de nouvelles croyances mais aussi de nouveaux contenus ; si, lorsque cela arrive, la mémoire fonctionne de manière fiable, elle est génératrice dans un sens plus fort [Michaelian, 2011a].

III. Justification des croyances mnésiques

Russell [1921] remarque la possibilité d'un scepticisme par rapport à notre connaissance du passé : comme il n'y a pas de connexion logique entre le fait que je semble me souvenir du passé et l'existence du passé, il est possible, par exemple, que le monde ait commencé à exister il y cinq minutes ; comment peut-on exclure cette possibilité ? Cet argument a suscité plusieurs réponses, mais la plupart des stratégies qu'on peut utiliser dans le but d'écarter l'argument sont comparables aux réponses standards faites au scepticisme concernant notre connaissance du monde extérieur. La question la plus intéressante pour l'épistémologie de la mémoire concerne donc les implications des épistémologies générales sur la justification des croyances mnésiques.

III.1. Fondationalisme

Si l'on suit la présentation de Senor [2009], on peut distinguer entre deux approches fondationalistes des croyances mnésiques : selon l'approche fondationaliste simple, toutes les croyances mnésiques sont justifiées par défaut ; selon l'approche fondationaliste expérientielle, une croyance mnésique est justifiée seulement lorsqu'elle est accompagnée par une image ou une apparence mnésique. Les épistémologues qui acceptent déjà le fondationalisme endosseront sans doute l'une ou l'autre de ces approches ; l'option rivale consiste à soutenir que le sujet est obligé d'établir sur d'autres fondations qu'il a le droit de se fier à sa mémoire. Mais les chances de réussite d'une telle stratégie semblent faibles [Alston, 1986] : comment pourrait-on établir qu'on a le droit de se fier à sa mémoire sans se fier à sa mémoire ?

Le fondationalisme simple et le fondationalisme expérientiel sont tous les deux problématiques. Premièrement, la thèse selon laquelle les croyances produites par la souvenance sont toujours justifiées par défaut est peu plausible, car la fiabilité de la souvenance dépend en partie de la fiabilité des autres processus de formation de croyances utilisés par le sujet

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

(ceci reste le cas même si l'on n'accepte pas un modèle de la photocopieuse primitif). Deuxièmement, l'approche fondationaliste ne s'accorde pas bien avec le rôle de la métacognition dans la formation des croyances mnésiques. Quand une représentation est récupérée de la mémoire, le sujet ne suppose pas automatiquement qu'elle est juste (même si, en termes phénoménologiques, il nous semble souvent que l'acceptation d'informations récupérées est automatique) ; au lieu d'une acceptation automatique, la surveillance métacognitive joue un rôle : se fiant à des indicateurs variés, le sujet décide d'accepter ou non la représentation. Le fondationaliste peut affirmer que, une fois que la surveillance métacognitive a conduit à l'acceptation de la représentation mnésique, la croyance résultante est toujours justifiée par défaut. Même cette position, cependant, ne s'accorde pas bien avec le rôle de la métacognition : la surveillance métacognitive joue un rôle dans la détermination de notre niveau de confiance dans nos propres croyances [Michaelian, 2012b], de sorte qu'un sujet peut former une croyance par la récupération mnésique mais peut néanmoins douter de sa justesse, ce qui suggère que les sujets ne se comportent pas toujours par rapport à leurs propres croyances comme si ces dernières étaient justifiées par défaut. Ces deux problèmes sont liés à un troisième, examiné par Senor [1993] dans le cas du fondationalisme simple : le fondationalisme simple, parce qu'il suppose que toutes les croyances mnésiques sont justifiées par défaut, implique qu'une croyance peut devenir justifiée simplement en étant stockée dans la mémoire, quel que soit son statut épistémique initial. Bien qu'il y ait des raisons indépendantes en faveur du fait de soutenir que la mémoire est capable d'engendrer de la justification [Lackey, 2005 ; Michaelian, 2011a], il est peu plausible de soutenir que la mémoire confère toujours de la justification, notamment dans les cas où la croyance était initialement injustifiée.

III.2. Cohérentisme

Comme le remarque Senor [2009], le cohérentisme négatif, selon lequel une croyance mnésique est justifiée pourvu qu'elle ne contredise pas les autres croyances du sujet, considère en fait les croyances mnésiques de la même manière que le fondationalisme simple : elles sont justifiées par défaut. Ainsi, le cohérentisme négatif est sujet aux mêmes problèmes que l'approche fondationaliste, examinés ci-dessus.

Le cohérentisme positif, selon lequel une croyance mnésique est justifiée seulement à condition qu'elle cohère de façon positive avec les autres croyances du sujet (où cela implique plus que le fait de ne pas contredire les autres croyances), se trouve exposé au problème des preuves

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

oubliées [Harman, 1986] : imaginez un cas où, quand un sujet acquiert une croyance, elle cohère de façon positive avec un grand nombre de ses autres croyances. Au fil du temps, ces autres croyances peuvent être oubliées, avec la conséquence que la nouvelle croyance ne cohère plus de façon positive. Pourvu que, entre-temps, le sujet n'ait pas acquis d'autres croyances avec lesquelles la nouvelle croyance puisse être incohérente, il semble que la croyance ne soit pas nécessairement devenue injustifiée. Mais ceci est la conclusion impliquée par le cohérentisme positif.

Pourvu qu'on s'intéresse toujours à une forme internaliste de la théorie, le cohérentisme positif et le cohérentisme négatif font tous les deux face à un problème supplémentaire, qu'on peut appeler le problème de la compartimentation. Selon le cohérentisme internaliste, une croyance est justifiée à condition non seulement qu'elle cohère avec les autres croyances du sujet mais également que le sujet soit (au moins en principe) capable de vérifier qu'elle cohère. Mais, comme la mémoire humaine est compartimentée, de telle façon qu'une croyance donnée se trouve normalement en contact uniquement avec un sous-ensemble relativement petit de l'ensemble global des croyances du sujet [Cherniak, 1986], le sujet ne devrait normalement pas être capable (même en principe) de vérifier qu'une croyance donnée cohère avec tout le reste de ses croyances.

III.3. Fiabilisme

Contrairement aux approches internalistes, le fiabilisme semble nous offrir une approche de la justification mnésique qui est défendable du point de vue empirique : à condition que la souvenance soit un processus de production de croyances fiable, les croyances mnésiques seront justifiées (bien entendu, cette justification peut être vaincue par des défaiseurs). Cette approche est trop simple, en revanche, même du point de vue du fiabiliste, car la mémoire semble être seulement un processus de production de croyances fiable sous conditions, soit un processus dépendant de croyances qui produit, de manière fiable, des croyances justes quand il est fourni avec des croyances justes [Goldman, 1979]. Si l'on voit la mémoire comme un processus dépendant de croyances, l'idée est que les croyances mnésiques sont justifiées parce que la mémoire conserve, de manière fiable, les croyances qui lui sont fournies par des processus fiables non dépendants de croyances, tels que la perception. Mais même cette approche reste trop simple, car elle ne prend pas en compte deux caractéristiques de la mémoire, déjà évoquées ci-dessus.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

Premièrement, l'endossement d'informations récupérées de la mémoire n'est pas automatique ; il est plutôt le résultat de la surveillance métacognitive du processus de récupération et des informations récupérées [Michaelian, 2012c]. Cela implique que la formation de croyances mnésiques peut être fiable même lorsque la plupart des croyances fournies à la mémoire sont fausses : si la surveillance métacognitive arrive à éliminer la plupart des représentations injustes récupérées de la mémoire, la formation de croyances mnésiques peut être fiable même lorsque la plupart des croyances fournies à la mémoire par d'autres sources de connaissances sont fausses. Il est possible que le contrôle de la source – dans le sens décrit par Mitchell et Johnson [2009] –, par exemple, joue un rôle important ici : selon la théorie du contrôle de la source, le sujet utilise des heuristiques qui dépendent des différences moyennes entre les souvenirs produits par des sources différentes afin d'évaluer la source probable (et donc la vérité probable) d'une représentation donnée ; dans la mesure où beaucoup de nos représentation mnésiques viennent de sources non fiables (l'imagination, les rêves, etc.), il est possible que le contrôle de source joue un rôle important dans l'explication de la fiabilité de la formation des croyances mnésiques. Le rôle de la métamémoire dans l'explication de la fiabilité de la souvenance a été malheureusement peu étudié en épistémologie, en dépit de l'existence d'un grand nombre de recherches empiriques et théoriques pertinentes provenant de la psychologie [Arango-Muñoz et Michaelian, 2014]. Beaucoup de ces recherches, par contre, ont été faites dans le domaine de la philosophie de l'esprit [Proust, 2013], même si son principal centre d'intérêt ne portait pas toujours sur la mémoire.

Deuxièmement, le point de vue naturel selon lequel la souvenance est un processus fiable à conditions (un processus dépendant de croyances qui produit, de manière fiable, des croyances justes quand il est fourni avec des croyances justes) présuppose implicitement le modèle de la photocopieuse, selon lequel la mémoire conserve mais ne produit pas de croyances. Si l'on prend en compte le caractère constructif de la mémoire, on risque de saper l'explication fiabiliste de la justification mnésique en prétendant que la souvenance n'est pas en fait un processus fiable à conditions. Un tel raisonnement a été proposé récemment par Shanton [2011], qui soutient que les erreurs mnésiques qui sont régulièrement produites par la mémoire constructive suggèrent que la souvenance n'est pas fiable. Une telle réaction à la recherche sur les erreurs mnésiques est certainement prématurée [Michaelian, 2013]. La recherche montre seulement que la mémoire n'est pas parfaitement fiable ; elle est faillible, ce qui est entièrement compatible avec le fait qu'elle soit suffisamment

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

fiable pour produire des croyances vraies. En outre, si l'on adopte une perspective adaptative sur la mémoire [Anderson, 1990], et si l'on suppose que les croyances vraies sont en général adaptatives et que les croyances fausses en général ne le sont pas, cela nous donne une raison évolutionnaire de supposer que les processus constructifs impliqués dans la mémoire produisent, en règle générale, des croyances vraies (même si la justesse peut être sacrifiée dans une certaine mesure dans le but d'obtenir un système qui soit adaptatif d'un point de vue global).

III.4. Épistémologie des vertus

Outre le fiabilisme, des approches externalistes intégrant une gamme de conditions modales, y compris plusieurs formes de l'épistémologie des vertus, ont gagné en influence ces dernières années. L'examen de deux versions de l'approche modale montrera qu'elles ont, semble-t-il, des implications différentes pour la connaissance mnésique. Comme Shanton [2011] le soutient, il semble que l'épistémologie des vertus développée par Ernest Sosa [2007], qui requiert pour la connaissance non seulement la fiabilité mais aussi la sécurité[1], implique qu'on a très peu de connaissances mnésiques épisodiques, car, étant donné le caractère reconstructif de la mémoire épisodique et les distorsions mnésiques occasionnelles qui en résultent, il arrive souvent que les souvenirs épisodiques ne satisfont pas la condition de sécurité, même lorsqu'ils sont vrais. Par ailleurs, il semble que l'approche anti-chance de Duncan Pritchard [2005] ne partage pas cette implication sceptique [Michaelian, 2013] : la condition anti-chance est beaucoup plus modeste que la condition de sécurité, et il est beaucoup plus probable qu'elle puisse être satisfaite par la mémoire épisodique, en dépit de son caractère constructif[2].

IV. Rôle des représentations mnésiques

IV.1. Réalisme direct

Selon le réalisme direct, soutenu notamment par Thomas Reid [1785], la mémoire est une forme de contact direct avec le passé. Une objection

1. La condition de sécurité est satisfaite en une occasion donnée si aucune croyance fausse n'est produite dans des mondes possibles qui en sont proches.

2. La condition anti-chance requiert seulement qu'il n'y ait pas de croyances fausses dans beaucoup de mondes possibles qui en sont proches.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

importante contre le réalisme direct est celle que Bernecker [2008] appelle l'objection de la cotemporalité : la thèse selon laquelle la mémoire ne peut pas nous mettre en contact direct avec des événements passés, car les événements n'existent pas au moment où l'on est censé en être conscient [Furlong, 1951]. Bernecker soutient que cette objection peut être évitée si l'on rejette le présentisme, thèse selon laquelle les événements passés n'existent pas.

L'argument de l'illusion, comparable à celui qu'on rencontre dans la philosophie de la perception [Furlong, 1951], constitue une autre objection importante : tout comme les illusions perceptuelles sont censées montrer que la perception ne nous met pas en contact direct avec la chose perçue, les faux souvenirs sont censés montrer que la mémoire ne nous met pas en contact direct avec la chose dont on se souvient. Bernecker suggère qu'il est possible d'éviter cette objection en adoptant une conception disjonctiviste de la mémoire selon laquelle le manque de contact, dans le cas de la mémoire illusoire, avec la chose dont on se souvient implique qu'il y a une différence métaphysique entre la mémoire authentique et la mémoire illusoire. Ainsi, ceux qui sont capables d'accepter la conception éternaliste des événements et la conception disjonctiviste des objets de la mémoire seront peut-être capables de soutenir le réalisme direct.

IV.2. Réalisme indirect

En termes métaphysiques, ces deux positions ne sont pas très attirantes, et le réalisme indirect, selon lequel notre contact avec le passé passe par des représentations mnésiques, reste la position la plus populaire. Comme le remarque Sutton [2010a], la critique la plus puissante du réalisme indirect, qui se base sur le simple constat que la mémoire ne semble pas normalement impliquer une conscience des représentations internes à partir desquelles l'événement passé est inféré, est évitée par une forme modérée de réalisme indirect, selon laquelle notre contact avec le passé peut passer par des représentations sans qu'il soit nécessaire d'être conscient de ces représentations elles-mêmes. Si l'on accepte une telle forme de réalisme indirect modéré, il s'ensuit que le sujet infère effectivement le passé, mais l'inférence est typiquement inconsciente.

Le réalisme indirect modéré est censé être une position conciliatoire, mais elle reste malgré tout controversée, comme la simple idée de traces mnésiques reste controversée. Sutton distingue plusieurs groupes d'objections à l'idée des traces ; on examinera ici seulement deux d'entre elles.

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

Une des objections formulées – qui provient de Ludwig Wittgenstein [1980] – est fondée sur l'idée selon laquelle, si la souvenance implique des traces, il faut que les traces pertinentes déterminent entièrement le contenu d'un souvenir. Comme le remarque Sutton, on peut rendre possible l'interaction des traces avec d'autres facteurs, y compris l'information disponible dans le contexte de récupération, afin de déterminer le contenu du souvenir. Une deuxième objection dépend du problème – relevé également par Wittgenstein [1980] – selon lequel des thèses philosophiques portant sur les traces mnésiques risquent d'être incompatibles avec les sciences empiriques de la mémoire. Les réponses à cette objection ont déjà été discutées ci-dessus, dans le contexte de la théorie causale de la mémoire.

Comme la théorie causale intègre l'idée des traces, on suppose normalement qu'elle implique le réalisme indirect. Bernecker [2008] soutient, cependant, qu'elle est compatible avec une forme de réalisme direct. C'est une bonne chose puisque, selon lui, le réalisme indirect requiert des marques de mémoire, des critères pour distinguer les représentations mnésiques véritables des représentations mnésiques seulement ostensibles (car, dans le cas contraire, il n'y aurait pas de base pour une inférence à l'événement passé [Holland, 1954]). Or nous ne disposons, selon lui, d'aucune théorie acceptable des marques de mémoire.

Hume [1739], par exemple, soutenait que les représentations mnésiques pouvaient être distinguées par leur relative flexibilité, ainsi que par leur vivacité (leur quantité de détails). Russell [1921] proposait comme marques de mémoire la quantité d'informations contextuelles, ainsi que les sentiments de passéité et de familiarité ; le sentiment de passéité a également été suggéré par un grand nombre d'autres théoriciens dont William James [1890] et C. D. Broad [1925]. Bernecker soutient que toutes ces propositions échouent, parce qu'elles n'isolent pas des marques fiables de mémoire, dans la mesure où la corrélation entre les marques et les véritables souvenirs est imparfaite. Mais le réaliste direct n'est pas obligé d'identifier des marques de mémoire qui soient en corrélation parfaite avec les véritables souvenirs, et il ne devrait pas tenter de le faire car, s'il y avait des marques parfaitement fiables, l'occurrence d'erreurs de mémoire serait inexplicable [Michaelian, 2012c].

En ce qui concerne la compatibilité de la théorie causale avec le réalisme direct, Bernecker [2008] admet que des traces du genre décrit par la théorie causale jouent un rôle en nous mettant en contact avec le passé : en vertu du fait d'avoir une représentation appropriée, le sujet est en contact avec l'événement passé, mais se souvenir de l'événement passé

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

n'exige pas d'être conscient de la représentation interne. Cette forme de réalisme direct modéré, cependant, semble être équivalente à la forme modérée de réalisme direct décrit par Sutton, et il est donc possible que le désaccord ici soit purement verbal.

V. Tendances émergentes

Ainsi, en qui concerne les questions centrales de l'épistémologie de la mémoire, l'état actuel du débat favorise les positions suivantes.

Théorie de la souvenance. Une version de la théorie causale, probablement modifiée de façon à la rendre compatible avec le caractère reconstructif de la souvenance, semble être l'option la plus attirante.

Justification des croyances mnésiques. Le fiabilisme est capable d'expliquer la justification des croyances mnésiques, mais il faut être prudent en ce qui concerne les détails ; certaines approches modales sont peut-être également défendables.

Rôle des représentations mnésiques. Un compromis entre le réalisme direct et le réalisme indirect semble être nécessaire afin d'éviter les problèmes rencontrés par les formes extrêmes de chaque position.

La présente section examine brièvement quelques tendances émergentes dans la recherche philosophique sur la mémoire vue comme source de connaissances. Le *memory boom* et l'explosion de la mémoire externe soulèvent des questions conceptuelles fondamentales sur la nature de la mémoire en tant que telle : la mémoire collective est-elle une forme de mémoire ? La mémoire externe est-elle une forme de mémoire ? Pourvu qu'on ait en tête la conception de la mémoire en termes de fidélité, qui met l'accent sur l'aspect constructif de la souvenance, il paraîtra relativement plausible que la mémoire collective, qui est produite par des compétitions et des négociations entre des groupes concurrents, et qui a souvent une correspondance assez faible avec le passé [Anastasio *et al.*, 2012], puisse être une forme de mémoire, mais relativement peu plausible que la mémoire externe, qui est conçue pour le stockage stable d'enregistrements discrets et immuables, sans prendre en compte la vérité ou la signification, puisse être une forme de mémoire. En outre, pourvu qu'on ait en tête la conception de la mémoire en termes de quantité, il paraîtra relativement plausible que la mémoire externe puisse

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

être une forme de mémoire, mais relativement peu plausible que la mémoire collective puisse être une forme de mémoire. Des réponses adéquates à ces questions devraient finalement intégrer les résultats de l'analyse philosophique de la mémoire.

V.1. La mémoire externe

Quoique la mémoire ait joué un rôle important dans les débats sur l'hypothèse de la cognition distribuée, elle a été utilisée la plupart du temps principalement comme exemple, et non comme objet d'investigation principal, et le débat philosophique a eu relativement peu de contact avec les recherches empiriques sur la mémoire externe et ses interactions avec la mémoire biologique. Cependant, un changement semble être en cours[1]. De plus, l'hypothèse de la cognition étendue a commencé à influencer la recherche empirique dans des champs tels que la psychologie [Fagin *et al.*, 2013] et l'archéologie [Jones, 2007] ; l'interaction est donc actuellement bidirectionnelle.

Il y a également un intérêt croissant pour les implications épistémologiques de l'hypothèse de la cognition étendue en général. Par exemple, Krist Vaesen [2011] a soutenu qu'il existe une tension entre l'épistémologie des vertus et l'hypothèse de la cognition étendue. Ce débat se déroule actuellement, et il est trop tôt pour pouvoir dire quel en sera le résultat[2], mais l'exemple central de la mémoire externe y jouera certainement un rôle important [Michaelian, 2014a].

V.2. La mémoire collective

Groupes à petite échelle. En rapport avec la recherche sur la cognition étendue et distribuée, il y a de plus en plus de recherches philosophiques sur la mémoire distribuée, ou inspirées par elles, dans les groupes à petite échelle. Par exemple, selon une position influente, le concept d'un système mnésique transactif décrit par Daniel Wegner [1987], qui a guidé des recherches empiriques dans plusieurs champs, n'implique pas une forme de mémoire collective qui est supérieure à la somme de ses parties [Pavitt, 2003]. Comme Theiner et d'autres chercheurs l'ont soutenu [Theiner *et al.,* 2010], pourtant, cette conclusion est sans

1. Un numéro spécial de la *Review of Philosophy and Psychology* est justement consacré à ce sujet ; voir Michaelian et Sutton [2013].

2. Voir Clark *et al.* [2012].

NE PAS DIFFUSER – LIVRE CONNAÎTRE – CHAP V
© ITHAQUE, 2015

doute prématurée, et des raisons convaincantes appuient l'idée que les systèmes mnésiques transactifs sont une forme émergente de mémoire de groupe.

Groupes à grande échelle. Actuellement, on constate un niveau de recherche très actif autour de la question de la mémoire collective. Quoique beaucoup de ces recherches, menées principalement par des historiens, soient de nature plutôt descriptive, Thomas Anastasio et ses collaborateurs [Anastasio *et al*., 2012] ont récemment essayé de passer de la description à l'explication, en développant un modèle unifié de la consolidation individuelle et collective. Il n'est pas certain que leur modèle puisse être appliqué de façon convaincante à la mémoire collective [Michaelian, 2014b], mais, avec un peu de chance, d'autres modèles plus convaincants seront bientôt proposés.

K. M.